Discussion Papers in Economics

Choice of models for emission-generating technologies and designing technical efficiency improvements.

Sushama Murty & Resham Nagpal

Discussion Paper 19-01

Centre for International Trade and Development

School of International Studies

Jawaharlal Nehru University

India

*

Choice of models for emission-generating technologies and designing technical efficiency improvements.

Sushama Murty[†*] and Resham Nagpal[††*]

January 5, 2019

[*]Centre for International Trade and Development, School of International Studies, Jawaharlal Nehru University, New Delhi 110067.

Emails: [†] sushama.murty1@gmail.com (Corresponding author) [††] reshamnagpal89@gmail.com

Choice of models for emission-generating technologies and designing technical efficiency improvements.

Abstract

Theoretical and empirical comparisons of weak-disposability(WD), joint-disposability(JD), by-production(BP), and input(I)-based approaches to modelling emission-generating technologies are made. Under constant or non-increasing returns, BP-based model satisfies JD, but stands unique in the class of DEA technologies, where graph-based indexes of technical efficiency improvements (TEIs) are identical for WD, JD, and I-based approaches, which are nested. Multi-relations-based BP approach is more successful in capturing the true data-generating process, and TEIs in input-usage result in an intuitive trade-off between optimal TEIs in good and bad output-production, not seen under other approaches. A table covering all configurations of optimal TEIs is derived for BP-approach and applied to study differences in optimal TEIs for non-performing production-units.

JEL classification codes: Q50, Q40, Q30, D24,

Keywords: modelling emission-generating technologies; free-input disposability of emission; weak disposability of good and bad outputs; joint disposability; by-production; graph versus output-based indexes of technical efficiency improvements; data envelopment analysis (DEA).

Choice of models for emission-generating technologies and designing technical efficiency improvements.

1 Introduction.

While the power sector has been a significant source of harmful emissions, accounting for nearly 40% of CO_2 emissions, 7% of primary $PM_{2.5}$, 48% of SO_2 emissions, and 28% of NO_x emissions world-wide in 2010, many recent interdisciplinary studies[1] have come to the conclusion that there is a potential for substantial emission reduction from addressing a disproportionately small share of global power plants also labeled as the "super-polluting units."[2] The studies show that opportunities to mitigate can be realised by a range of operational changes in the functioning of super-polluting units such as installing pollution control technologies, using higher quality coal, replacement of these units by units with higher electric efficiency, or simply by retiring non-performing units.

In economics, a parallel approach for identifying and targeting non-performing or under-performing production units is the study and measurement of technical efficiency. Production units that lie far below the production frontier defined by the most efficient units in the dataset are deemed as non-performing or under-performing. While the primary focus of the inter-disciplinary literature mentioned above is to identify super-polluting units and correlate their emission volumes to their qualitative features, a technical efficiency index of a producing unit was envisioned as a more holistic concept that focuses not just on its performance in the emission dimension but aggregates its performance along various dimensions including production of the desirable output (here electricity generation), generation of undesirable emission, and usage of costly inputs. The computation of a technical efficiency index hence also involves an important normative component, namely, a choice of policy weights to be assigned to each of the different dimensions of performance. These determine the extent to which generation of harmful emissions has to be discounted, and can significantly influence the performance ranking of units.

The computation of a technical efficiency index identifies, for every producing unit, a configuration of optimal technical efficiency improvements in production of the outputs and/or the usage of the inputs that places the producing unit on the production frontier defined by the most efficient units in the dataset under some standard economic assumptions. Such technical efficiency improvements can be conceptualised as proportional increases in the production of the

[1]See for example, Tong et al (2018) and references there in.

[2]For example, according to Tong et al (2018), 14.2% of global primary $PM_{2.5}$ emissions from coal-fired plants were produced by just 0.7% of total capacity.

desirable/intended/good output and proportional reductions in the usage of inputs and generation of the undesirable emission/bad output. Hence, a technical efficiency index can be shown to map into an index of technical efficiency improvements. Such indexes are output-based if input usage is held fixed and efficiency improvements are allowed only in the directions of the outputs (both good and bad). They are graph-based if efficiency improvements are permitted in all input and output dimensions.

A pioneering and highly influential work in the area of measuring technical efficiency of producing units generating bads is Färe, Lovell, and Pasurka (1989). Since then, there has been a slew of studies that have applied the methodology proposed in this paper.[3] At the same time, several critiques of this approach and other alternative, and sometimes rival, methodologies have also developed in the literature.[4] These methodologies differ only with the respect to the approach taken to model an emission-generating technology relative to whose efficient frontier technical efficiency is to be measured. Extending the model of a standard neo-classical technology to incorporate production of bad outputs is challenging. This is because emission is an example of an output that is not freely disposable and empirical observations show that its generation is positively related with the production of the good output. It is hence tempting to treat it as an ordinary input. But standard production theory (where free disposability of all inputs is assumed) implies a non-positive relation between any two inputs along the isoquant. Hence, if emission is treated as an ordinary input, this would imply a non-positive trade-off between the emission and any emission-generating input such as fossil fuel along the efficient frontier of the technology, which is counterintuitive.[5]

In this work we theoretically and empirically compare four influential technology modelling approaches and their implications for measurement of technical efficiency and the design of optimal technical efficiency improvements. These are the weak-disposability based output approach adopted in the original Färe, Lovell, and Pasurka (1989) work; the classic input approach attributable to Baumol and Oates (1975, 1988) and Cropper and Oates (1992); and the more recent by-production approach of Murty, Russell, and Levkoff (2012) (henceforth MRL) and the joint disposability-based approach of Ray, Mukherjee, and Venkatesh (2016/2018).

These technology-modelling approaches differ with respect to the disposability conditions assumed for the good and bad outputs and the emission-causing inputs. However, we show that, under constant or non-increasing returns to scale, the input approach satisfies weak-

[3]See for example, Boyd and McClelland (1999), Coggins and Swinton (1996), Murty and Kumar (2002, 2003), and Hailu and Veeman (2000). Zhou et al (2008) reviews over a hundred papers that have applied this methodology.

[4]See for instance Førsund (1998, 2009, and 2017), Murty and Russell (2002) and Murty, Russell, and Levkoff (2012) for critiques of this methodology and Murty and Russell (2017) for a general survey of recent developments in this area.

[5]See Murty, Russell, and Levkoff (2012), Murty and Russell (2017), and Murty (2015).

disposability of the good and bad outputs as defined in Färe et al (1989), while the by-production approach satisfies joint-disposability of the emission-causing inputs and the good and bad outputs.[6]

The data envelopment approach (DEA) is one of the most popular methodologies employed to study and measure technical efficiency. DEA constructs technologies with piece-wise linear frontiers based on modeller-specified disposability conditions and restrictions such as convexity and returns to scale. In the class DEA technologies, we find a clear nesting: the sub-class of technologies based on the input approach nests the sub-class of technologies based on the weak disposability-based output approach, which in turn nests the sub-class of technologies based on the joint disposability approach. Moreover, all these three types of technologies, when constructed from the same dataset, share the same strictly efficient frontier, although their weakly efficient frontiers may differ. This implies that the values of the weighted graph indexes of efficiency improvements and the optimal efficiency improvements they entails are the same for these three technological specifications. In addition, the strictly efficient frontiers of the production possibility sets, which are defined in the space of good and bad outputs when usage of all inputs are held fixed, coincide for the input and weak disposability-based approaches. This implies that the values of the output-based weighted index of efficiency improvements are also same for the input and weak disposability-based approaches, but they may differ from the value of the index derived for a jointly disposable technology.

The paper highlights the importance of evaluating the soundness of technology modelling approaches on the basis of the extent to which they are capable of capturing features of the true data-generating processes. The dataset of the Indian Central Electricity Authority (CEA) on coal-based thermal power sector, where data on emission is generated by a linear formula employing constant emission factors of coal and oil, provides a unique test for the modelling approaches. The tiny share of oil implies that the relation between CEA data on aggregate fossil fuel input (measured in heat units) and emission generation is nearly linear, with the reference constant of proportionality being the emission factor of coal. However, we find that the ratio of reduction in the *efficient* level of emission generated and the reduction in the fossil fuel input varies significantly across plants for the weak-disposability and joint-disposability based approaches, while it is constant under the by-production approach and close to the emission factor of coal. Moreover, in the case of the former two approaches, there are plants for which optimal proportional reduction in emission is zero, even when there is a positive reduction in usage of fossil-fuel input, which seems counterintuitive given that fossil fuels have positive emission-factors.

[6]This is an extension of original joint disposability assumption of Ray et al (2016/2018). See also Section 2.2 of this paper.

A decomposition of the optimisation that computes the weighted graph index of efficiency improvements shows that, under all technological specifications studied in this paper, optimal efficiency improvements in good and bad output production can be expressed as functions of efficiency improvements in input usage. In particular, under the by-production approach, there is a trade-off between optimal efficiency improvements in production of the good and bad outputs: the former is non-increasing, while the latter is non-decreasing in the extent of efficiency improvements in inputs.

We tabulate all possible configurations of optimal efficiency improvements in inputs and outputs for a by-production technology. These possible solutions of the optimisation that computes the weighted graph index of efficiency improvements reflect the trade-off discussed above. In particular, starting from the initial production vector, when there is positive proportional reductions in emission-causing inputs, then the maximum possible reductions in emission do not fall but can increase, while the maximum possible increase in the good output production cannot increase but can fall. This non-positive correlation may not be displayed by solutions obtained under other technological modelling approaches, as monotonicity of at most the weighted average of optimal efficiency improvements in good and bad output production can be ensured for these approaches. This is consistent with the efficient level of emission increasing under these modelling approaches even when there are positive proportional decreases in usage of emission-causing inputs, which is counterintuitive. The table of all possible optimal configurations of efficiency improvements under the by-production approach is applied to study the differences in optimal efficiency improvements across power plants in our dataset.

Section 2 reviews the four technology modelling approaches studied in this paper and explores the theoretical relations between them. Section 3 defines the weighted graph and output-based indexes of efficiency improvements and studies the implications of relations established in Section 2 for measurement of these indexes. These indexes are also employed to measure efficiency improvements in good and bad outputs that can be attributed purely to efficiency improvements in usage of inputs. Section 4 shows a decomposition of the optimisation problem defining the weighted graph index of efficiency improvements that yields efficiency improvements in production of good and bad outputs as functions of efficiency improvements in inputs. The implications for by-production technology modelling are studied and the monotonicity properties of efficiency improvements in production of good and bad outputs under different technological specifications are established. Section 5 discusses the special features of the CEA dataset, while Section 6 provides an empirical comparison of different modelling approaches and the optimal efficiency improvements they entail. Section 7 tabulates all possible configurations of optimal efficiency improvements under the by-production approach in Table 5 and applies this table to the CEA data set on Indian coal-based thermal power sector. Section 8 concludes.

2 Alternative specifications of emission-generating technologies.

We assume that there is one marketed (or good/intended/economic) output, one bad output, and n inputs. A production vector is denoted by $\langle x, y, z \rangle \in \mathbf{R}^{n+2}_{+}$, where $x \in \mathbf{R}^{n}_{+}$ denotes the vector of quantities of the n inputs, y is the net amount of the intended output produced, and z is the quantity of emission generated.

2.1 A by-production approach.

In this paper, we study the case where a by-production technology is obtained as an intersection of two sub-technologies.[7] The first, denoted by $T_1 \subset \mathbf{R}^{n}_{+}$, is a standard neo-classical technology based on human engineering design that describes the transformation of all inputs into the good output. Some of the inputs used in intended production are composed of emission-causing substances in proportions determined by nature. When such inputs are employed in intended production, relations defined by the second sub-technology $T_2 \subset \mathbf{R}^{n}_{+}$ based on considerations such as the laws of thermodynamics become operational and emission is generated from emission-causing substances and their reactions with other substances in the production environment.[8]

To distinguish between emission-causing and non-emission causing inputs, we partition the vector of input quantities into $x = \langle x_o, x_z \rangle \in \mathbf{R}^{n}_{+}$, where $x_o \in \mathbf{R}^{n_o}_{+}$ and $x_z \in \mathbf{R}^{n_z}_{+}$ denote, respectively, the vectors of quantities of non-emission causing and emission-causing inputs used and $n_o + n_z = n$. For $i = 1, \ldots, n_o$, the i^{th} non-emission causing input quantity is denoted by x_{o_i}, while for $i = 1, \ldots, n_z$, the i^{th} emission-causing input quantity is denoted by x_{z_i}. Depending on convenience, we sometimes also index inputs simply by $i = 1, \ldots, n$, so that x_i denoted the amount of the i^{th} input.

To facilitate a formal definition of a by-production technology, we first define the *costly*

[7] In its most general formulation, a by-production technology is obtained as an intersection of two or more sub-technologies, each of which captures a distinct rule governing the transformation of inputs into the good and bad outputs. See MRL, Murty(2015), and Murty and Russell (2016/2018, 2017). See Serra et al (2016) for an empirical application. A multi-relation formulation for defining some types of production processes was first proposed by Nobel laureate Ragnar Frisch (1965).

[8] For example, during production of thermal electricity from coal, engineers are concerned with the gross calorific value (GCV) of coal, which measures its heat content, as it is heat energy that is ultimately transformed into thermal electricity in power plants. However, because coal contains the emission-causing substance carbon, its use for electricity generation also generates CO_2 emission. The extent of emission generated depends on the emission-factor of the coal-type employed (which measures the carbon content per unit of coal) and the exposure of the power plant to oxygen.

disposal hull of any set $A \subset \mathbf{R}_+^{n+2}$ as the set[9]

$$CDH(A) \;\; := \;\; \Big\{ \langle x_o, x_z, y, z \rangle \in \mathbf{R}_+^{n+2} \; \Big| \; \exists \; \langle x_o, x'_z, y, z' \rangle \in A \text{ with } x'_z \geq x_z \text{ and } z' \leq z \Big\}.$$

From its definition, the costly disposal hull $CDH(A)$ includes set A and all production vectors that contain arbitrarily larger amounts of the emission and arbitrarily lower amounts of the emission-causing inputs than those in set A.

Definition 1 *A set $T \subset \mathbf{R}_+^{n+2}$ is a by-production technology (BPT) if there exist two closed sets $T_1 \subset \mathbf{R}_+^{n+2}$ and $T_2 \subset \mathbf{R}_+^{n+2}$ such that the following hold:*[10]

- $T = T_1 \cap T_2$
- *Set T_1 satisfies*

  (i) *free disposability of the intended output and inputs:*

$$\langle x, y, z \rangle \in T_1 \;\wedge\; \bar{y} \leq y \;\wedge\; \bar{x} \geq x \Longrightarrow \langle \bar{x}, \bar{y}, z \rangle \in T_1. \tag{1}$$

  (ii) *independence from emission generation:*

$$\langle x_o, x_z, y, z \rangle \in T_1 \;\wedge\; \bar{z} \neq z \;\Longrightarrow\; \langle x_o, x_z, y, \bar{z} \rangle \in T_1.$$

  (iii) *convexity.*

- *Set T_2 satisfies*

  (i') *independence from production of the good output and usage of non-emission causing inputs:*

$$\langle x_o, x_z, y, z \rangle \in T_2 \;\wedge\; \bar{y} \neq y \;\wedge\; \bar{x}_o \neq x_o \;\Longrightarrow\; \langle \bar{x}_o, x_z, \bar{y}, z \rangle \in T_2.$$

[9] Vector notation: Given two vectors a and b in $\mathbf{R}^n$,

$$\begin{aligned} a &\geq b \iff a_i \geq b_i \;\forall\; i = 1, \ldots, n \\ a &> b \iff a \neq b \text{ and } a_i \geq b_i \;\forall\; i = 1, \ldots, n \\ a &\gg b \iff a_i > b_i \;\forall\; i = 1, \ldots, n \end{aligned}$$

[10] This definition of a by-production technology is weaker than the one in MRL, Murty (2015), and Murty and Russell (2016/2018).

(ii')

$$
\begin{aligned}
G\left(x_o, x_z, y; T_2\right) \;&:=\; \inf\left\{z \geq 0 \mid \langle x_o, x_z, y, z\rangle \in T_2\right\} \qquad (2)\\
&=\; \inf\left\{z \geq 0 \mid \langle x_o, x_z, y, z\rangle \in CDH(T_2)\right\} \;=:\; G\left(x_o, x_z, y; CDH(T_2)\right).
\end{aligned}
$$

(iii') *convexity.*

Conditions (i) and (iii) in the definition above include, respectively, standard disposability and convexity assumptions that are usually imposed on neo-classical technologies.[11] Condition (ii) implies that sub-technology T_1 imposes no restrictions on the level of emission. If a production vector belongs to T_1, then so does any other production vector with the same amounts of inputs and the intended outputs but with *any other* amount of the emission. Thus, it is assumed that the production of the intended output is unaffected by the level of emission.[12]

Condition (i') says that emission generation is not directly caused by the intended output or the non-emission causing inputs.[13] Changes in the levels of production or usage of these goods do not affect emission-generation if the levels of the emission-causing inputs are held fixed.

In condition (ii'), functions $G\left(x_o, x_z, y; T_2\right)$ and $G\left(x_o, x_z, y; CDH(T_2)\right)$ define the lower frontiers of sub-technology T_2 and its costly disposal hull, respectively. They give the minimum amounts of emission that can be generated under sub-technology T_2 and its costly disposal hull $CDH(T_2)$, respectively, when levels of all other other goods, including the emission-causing inputs, are held fixed.[14] Condition (ii'), which has been discussed at length in Murty and Nagpal (2018), requires that the lower frontiers of sub-technology T_2 and its costly disposal hull coincide. Murty and Russell (2017) show that the lower frontier of $CDH(T_2)$ is always non-negatively sloped.[15] Hence, condition (ii') captures most real-life cases where the minimum emission generated during production increases with increase in use of emission-causing inputs, *e.g.* CO_2 emission level is increasing in the usage of coal. Figure 1 assumes that $n = n_z = 1$.[16] While the upper panels (a) and (c) are examples of set T_2, the lower panels (b) and (d) are

[11] See, for instance, Shephard (1953), Debreu (1959), and Mas-Colell et al (1995) Chapter 5.

[12] For a treatment of a more general case under the by-production approach, where it can be, see Murty (2015).

[13] For a generalisation, where emissions can also be caused by intended outputs, see Murty (2015).

[14] Condition (i') implies that the minimum levels of emission generated are independent of the levels of the intended output and non-emission causing inputs. Thus, we can re-write these functions as $G\left(x_z; T_2\right)$ and $G\left(x_z; CDH(T_2)\right)$.

[15] This is because $CDH(T_2)$ satisfies the following costly disposability assumptions defined in MRL and Murty and Russell (2017): $\langle x_o, x_z, y, z\rangle \in CDH(T_2) \;\wedge\; \bar{z} \geq z \;\wedge\; \bar{x}_z \leq x_z \implies \langle x_o, \bar{x}_z, y, \bar{z}\rangle \in CDH(T_2)$, which Murty and Russell (2017) have shown to imply a non-negative relation between emission-causing inputs and emission along the lower frontier of $CDH(T_2)$. In other words, $G\left(x_o, x_z, y, CDH(T_2)\right)$ is non-decreasing in x_z.

[16] Figure 1 also assumes that condition (i') is true, so that set T_2 and its costly disposal hull can be effectively studied in the space formed by the axes of the emission and the single-input.

their respective costly disposal hulls. Condition (ii') rules out cases such as those depicted in panels (c) and (d), where the lower frontiers of T_2 and its costly disposal hull do not coincide. In particular, the lower frontier of sub-technology T_2 has a downward sloping region, indicated by the dashed curve in panel (d), which is not a part of the lower frontier of the costly disposal hull of sub-technology T_2. The latter, as seen, is non-negatively sloped.

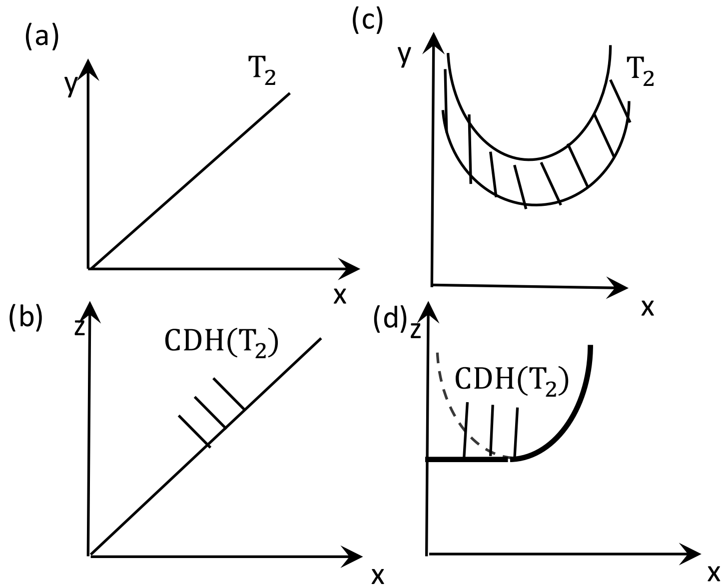

Figure 1

Since this work studies optimal technical efficiency improvements, our focus is on measuring how far are the operations of producing units from the efficient frontier of technology T. In the case of a BPT, this boils down to measuring how far are the production vectors of these units from the efficient frontier of sub-technology T_1 and the lower frontier of sub-technology T_2. Since under condition (ii') of Definition 1 of a BPT, both sets T_2 and $CDH(T_2)$ share the same lower frontier, it suffices to focus on the latter while computing technical efficiency improvements. Thus we estimate the BPT, $T^{BP} := T_1 \cap CDH(T_2)$, in lieu of the original BPT, $T_1 \cap T_2$.

To obtain a DEA representation of a BPT, suppose producing units are indexed by $u = 1, \ldots, U$. The $U \times n$-dimensional data matrix of inputs is denoted by

$$X = \begin{bmatrix} X_o & X_z \end{bmatrix},$$

where X_o and X_z are, respectively, the $U \times n_o$ and $U \times n_z$-dimensional data matrices of non-emission causing and emission-causing inputs. The $U \times 1$-dimensional data matrices of intended output and emission are denoted by Y and Z, respectively. Then, assuming technological constant returns to scale, the DEA representation of T^{BP} follows from MRL as[17]

$$T^{BP} = \Big\{ \langle x_o, x_z, y, z \rangle \in \mathbf{R}_+^{n+2} \ \Big| \ \lambda^\top X \leq x^\top, \quad \lambda^\top Y \geq y, \quad \mu^\top X_z \geq x_z^\top, \quad \mu^\top Z \leq z, \quad \lambda \geq 0_U, \quad \mu \geq 0_U \Big\}. \tag{3}$$

2.2 The weak and joint disposability-based output and input approaches.

Given a technology T, define the production-possibility set corresponding to any vector of inputs $x \in \mathbf{R}_+^n$ as the set of all good and bad output combinations that are producible by input vector x:

$$P(x) := \{\langle y, z \rangle \in \mathbf{R}_+^2 \mid \langle x, y, z \rangle \in T\}.$$

Similarly, for any vector of non-emission causing inputs $x_o \in \mathbf{R}_+^{n_o}$, we can also define the set

$$P(x_o) := \{\langle x_z, y, z \rangle \in \mathbf{R}_+^{n_z+2} \mid \langle x_o, x_z, y, z \rangle \in T\}.$$

2.2.1 Technological specifications under weak and joint disposability-based output and input approaches.

The weak disposability-based output approach of Färe et al (1989) is predicated on the assumptions of null-jointness (no emission-generation implies there has been no production of the intended output)

$$\langle y, z \rangle \in P(x) \quad \wedge \quad z = 0 \implies y = 0$$

and weak disposability (*i.e.,* simultaneous radial disposability) of emission and the intended output:

$$\langle y, z \rangle \in P(x) \implies \langle \lambda y, \lambda z \rangle \in P(x) \quad \forall \ x \in \mathbf{R}_+^n \text{ and } \lambda \in [0, 1].$$

While it makes no further disposability assumptions on emission, it assumes standard free disposability of all inputs and intended outputs as stated in (1). The following DEA model of a technology first proposed by Färe et al (1989) satisfies weak disposability, free disposability

[17] Where the DEA representations of sets T_1 and $CDH(T_2)$ are

$$\begin{aligned} T_1 &= \Big\{ \langle x, y, z \rangle \in \mathbf{R}_+^{n+2} \ \Big| \ \lambda^\top X \leq x^\top, \ \lambda^\top Y \geq y, \ \lambda \geq 0_U \Big\} \text{ and} \\ CDH\left(T_2\right) &= \Big\{ \langle x_o, x_z, y, z \rangle \in \mathbf{R}_+^{n+2} \ \Big| \ \mu^\top X_z \geq x_z^\top, \ \mu^\top Z \leq z, \ \mu \geq 0_U \Big\}. \end{aligned}$$

of the good output, free disposability of all inputs, constant returns to scale, and convexity:

$$T^{WD} \ := \ \Big\{ \langle x, y, z \rangle \in \mathbf{R}_{+}^{n+2} \ \Big| \ \lambda^{\top} X \leq x^{\top}, \ \lambda^{\top} Y \geq y, \ \lambda^{\top} Z = z, \ \lambda \geq 0_U \Big\}, \tag{4}$$

The original joint disposability approach of Ray et al (2016/2018) assumes weak disposability (*i.e.,* simultaneous radial disposability) of emission-causing inputs and the emission, while the intended output and the remaining inputs are assumed to satisfy standard free disposability assumptions as defined in (1). Murty and Russell (2017) argue that in the Ray et al (2018) model, radial reductions in usage of emission-causing inputs and emission generation do not have any adverse impact on production of the good output. This could be considered counterintuitive as it may not be possible to maintain existing levels of good-output production when emission-causing inputs are reduced holding all other inputs fixed. The extension below of the original joint disposability approach of Ray et al (2018) assumes weak disposability of emission-causing inputs, emission, and the intended output:[18]

$$\langle x_z, y, z \rangle \in P(x_o) \implies \langle \lambda x_z, \lambda y, \lambda z \rangle \in P(x_o) \qquad \forall \ x_o \in \mathbf{R}_{+}^{n_o} \text{ and } \lambda \in [0, 1].$$

We will be employing this definition of joint disposability in this paper. The following DEA model of a technology satisfies jointly disposability, constant returns to scale, and convexity

$$T^{JD} \ := \ \Big\{ \langle x, y, z \rangle \in \mathbf{R}_{+}^{n+2} \ \Big| \ \lambda^{\top} X_o \leq x_o^{\top}, \ \lambda^{\top} X_z = x_z^{\top}, \ \lambda^{\top} Y = y, \ \lambda^{\top} Z = z, \ \lambda \geq 0_U \Big\}, \tag{5}$$

In contrast, in addition to technology satisfying standard free disposability with respect to all inputs and the intended output, the input approach treats emission as a standard input satisfying input free disposability:

$$\langle x, y, z \rangle \in T \ \wedge \ \bar{z} \geq z \implies \langle x, y, \bar{z} \rangle \in T,$$

The following DEA model of a technology satisfies free-input disposability of emission, constant returns to scale, and convexity

$$T^{I} \ = \ \Big\{ \langle x, y, z \rangle \in \mathbf{R}_{+}^{n+2} \ \Big| \ \lambda^{\top} X \leq x^{\top}, \ \lambda^{\top} Y \geq y, \ , \ \lambda^{\top} Z \leq z, \ \lambda \geq 0_U \Big\}. \tag{6}$$

2.2.2 Some theoretical relations between different technological specifications.

In empirical work, technologies are usually assumed to satisfy some returns to scale assumption. A consequence of this is summarised in the remark below:

[18] We are grateful to the authors for bringing this extension to our notice.

Proposition 2 *Suppose T is an emission-generating technology satisfying non-increasing or constant returns to scale.*

(i) If T satisfies free disposability of all inputs, then it satisfies weak disposability of intended output and emission.

(ii) If T satisfies free disposability of all non-emission causing inputs, then it satisfies joint disposability of emission-causing inputs, intended output, and emission.

Proof. (i) Suppose $\langle y, z\rangle \in P(x)$ and $\lambda \in [0, 1]$. Non-increasing or constant returns to scale and free disposability of all inputs implies $\langle \lambda y, \lambda z\rangle \in P(\lambda x) \subseteq P(x)$. Hence, $\langle \lambda y, \lambda z\rangle \in P(x)$.
(ii) Suppose $\langle x_z, y, z\rangle \in P(x_o)$ and $\lambda \in [0, 1]$. Non-increasing or constant returns to scale and free disposability of non-emission causing inputs implies $\langle \lambda x_z, \lambda y, \lambda z\rangle \in P(\lambda x_o) \subseteq P(x_o)$. Hence, $\langle \lambda x_z, \lambda y, \lambda z\rangle \in P(x_o)$. ■

Remark 3 *It follows from Theorem 2 that, under non-increasing or constant returns to scale, the input approach also implies weak-disposability of emission and the intended output, while the BPT defined in MRL is also jointly disposable under non-increasing or constant returns to scale as it assumes free disposability of non-emission causing inputs.*

From the definitions of DEA technologies T^{WD}, T^{JD}, and T^{I} defined in (4), (5), and (6), respectively, it follows that[19]

$$T^{JD} \subset T^{WD} \subset T^{I}. \tag{7}$$

Though the weakly efficient frontiers of these technologies differ, the theorem below states that they all share the same strictly efficient frontier.[20] The result follows the fact that if $\langle x_o, x_z, y, z\rangle$ is a strictly efficient production vector of T^{I}, then there is no wastage of inputs, no under-production of the good output, or no over-production of the bad output. Hence, all inequalities in (6) hold as equalities. It then follows from (4) and (5) that $\langle x_o, x_z, y, z\rangle$ is in T^{WD} and T^{JD}. Further, (7) implies that $\langle x_o, x_z, y, z\rangle$ is also a strictly efficient production vector of T^{WD} and T^{JD}. The theorem also states that the strictly efficient frontiers of the production possibility sets of T^{WD} and T^{I} corresponding to any given vector of inputs coincide.

[19] For example, if $\langle x_o, x_z, y, z\rangle$ satisfies inequalities (5), then in also satisfies inequalities in (4) and (6).

[20] A production vector $\langle x_o, x_z, y, z\rangle$ in set T is a *strictly efficient point* of T if there exists no other point in T with no bigger amounts of the inputs and emission and no smaller amount of the good output, *i.e.,* if there exists no other point $\langle x'_o, x'_z, y', z'\rangle$ in T such that $\langle x'_o, x'_z\rangle \leq \langle x_o, x_z\rangle$, $y' \geq y$, and $z' \leq z$.
A production vector $\langle x_o, x_z, y, z\rangle$ in set T is a *weakly efficient point* of T if there exists no other point in T with smaller amounts of the inputs and emission and bigger amount of the good output, *i.e.,* if there exists no other point $\langle x'_o, x'_z, y', z'\rangle$ in T such that $\langle x'_o, x'_z\rangle \ll \langle x_o, x_z\rangle$, $y' > y$, and $z' < z$.

Theorem 4 *(i) The strictly efficient frontiers of DEA technologies T^{WD}, T^{I}, and T^{JD} coincide.*
(ii) Define the production possibility sets of T^{WD} and T^{I} for any vector of inputs $x \in \mathbf{R}^n_+$:

$$
\begin{aligned}
P^{I}(x) &= \{\langle y, z\rangle \in \mathbf{R}^2_+ \mid \lambda^\top X \leq x^\top,\ \lambda^\top Y \geq y,\ \lambda^\top z \leq z,\ \lambda \geq 0_U\} \quad \text{and} && (8)\\
P^{WD}(x) &= \{\langle y, z\rangle \in \mathbf{R}^2_+ \mid \lambda^\top X \leq x^\top,\ \lambda^\top Y \geq y,\ \lambda^\top z = z,\ \lambda \geq 0_U\}, && (9)
\end{aligned}
$$

respectively. Then the strictly efficient frontiers of sets $P^{WD}(x)$ and $P^{I}(x)$ also coincide.[21]

3 Weighted indexes of efficiency improvements.

To define a weighted index of efficiency improvements, we first define the set of all non-negative weights on inputs and outputs which sum to one as the unit simplex

$$\Delta = \Big\{ w = \langle w_o^x, w_z^x, w^y, w^z \rangle \in \mathbf{R}^{n+2}_+ \ \Big|\ \sum_{i=1}^{n_o} w^x_{o_i} + \sum_{i=1}^{n_z} w^x_{z_i} + w^y + w^z = 1 \Big\},$$

where $w_o^x \in \mathbf{R}^{n_o}_+$, $w_z^x \in \mathbf{R}^{n_z}_+$, w^y, and w^z denote the weights on non-emission and emission-causing inputs and the good and bad outputs, respectively.

3.1 A weighted index of graph efficiency improvements.

As in MRL, let $\otimes$ denote an operator in any m-dimensional Euclidean space $\mathbf{R}^m_+$ such that, for any two vectors a and b in $\mathbf{R}^m_+$, we have

$$a \otimes b \;=\; \langle a_1 b_1,\ \ldots\ ,\ a_m b_m \rangle.$$

Starting from an initially given production vector $v := \langle x_o, x_z, y, z\rangle$ in technology $T \subset \mathbf{R}^{n+2}_+$, let $\delta_o \in \mathbf{R}^{n_o}$ and $\delta_z \in \mathbf{R}^{n_z}$ denote proportional reductions in the usage of the non-emission causing and the emission-causing inputs, respectively, and let scalars θ and γ denote proportional increase and proportional decrease in the production of the intended output and the emission, respectively.

Definition 5 *The vector of proportional changes in inputs and outputs $\langle \delta, \theta, \gamma\rangle = \langle \delta_o, \delta_z, \theta, \gamma\rangle \in \mathbf{R}^{n+2}_+$ is a vector of efficiency improvements if*

(i) $\Big\langle x_o - (\delta_o \otimes x_o)\,,\quad x_z - (\delta_z \otimes x_z)\,,\quad y + \theta y,\quad z - \gamma z \Big\rangle \in T$ *and*

[21] Given a technology T, a vector $\langle y, z\rangle$ is a strictly efficient point of $P(x)$ if it lies in set $P(x)$ and there exists no other point $\langle \bar{y}, \bar{z}\rangle \in P(x)$ such that $\bar{y} \geq y$ and $\bar{z} \leq z$.

(ii) $\Big\langle x_o - (\delta_o \otimes x_o)\,,\ x_z - (\delta_z \otimes x_z)\Big\rangle \leq \langle x_o, x_z\rangle$; $z - \gamma z \leq z$; *and* $y + \theta y \geq y$ *with at least one of the weak inequalities holding as a strict inequality.*

In the above definition, (i) implies that the vector of proportional changes in inputs and outputs $\langle \delta, \theta, \gamma\rangle$ is feasible under technology T and (ii) implies that it results in no-bigger amounts of inputs and emission and no lower amount of the intended output. Note that non-negativity of a feasible production vector imply that $\delta_o \in [0,1]^{n_o}$, $\delta_z \in [0,1]^{n_z}$, $\gamma \in [0,1]$, and $\theta \geq 0$. We define the weighted index of graph efficiency-improvements for technology $T \in \{T^{BP}, T^{WD}, T^{JD}, T^{I}\}$ as a mapping $\mathfrak{I}^G_T : T \times \Delta \longrightarrow \mathbf{R}_+$ with image[22]

$$\begin{aligned} \mathfrak{I}^G_T(x_o, x_z, y, z; w) = \max_{\langle \delta_o, \delta_z, \theta, \gamma\rangle} \ &\sum_{i=1}^{n_o} w^x_{o_i}\delta_{o_i} + \sum_{i=1}^{n_z} w^x_{z_i}\delta_{z_i} + w^y\theta + w^z\gamma \\ &\text{subject to} \\ \Big\langle x_o - (\delta_o \otimes x_o)\,,\quad x_z - (\delta_z \otimes x_z)\,,\quad &y + \theta y, \quad z - \gamma z \Big\rangle \in T, \\ \delta_o \in [0,1]^{n_o}, \quad \delta_z \in [0,1]^{n_z}, \quad &\gamma \in [0,1], \quad \theta \geq 0. \end{aligned} \tag{10}$$

Starting from any given production vector $v = \langle x_o, x_z, y, z\rangle \in T$ and a vector of weights $w \in \Delta$, Problem (10) finds the vector of efficiency improvements that maximises the weighted sum of proportional reductions in all inputs and the bad output and proportional increase in the good output. It is clear that if a vector $s^*_T := \langle \delta^*_{oT}, \delta^*_{zT}, \theta^*_T, \gamma^*_T\rangle \in \mathbf{R}^{n+2}_+$ is a solution to Problem (10) for technology T, then $v_{s^*_T} := \Big\langle x_o - (\delta^*_{oT} \otimes x_o)\,,\quad x_z - (\delta^*_{zT} \otimes x_z)\,,\quad y + \theta^*_T y, \quad z - \gamma^*_T z\Big\rangle$ is a strictly efficient point of technology T.[23] The remark below follows as DEA technologies T^{WD}, T^{JD}, and T^I have the same strictly efficient frontiers.

Remark 6 *Given any data point* $v = \langle x_o, x_z, y, z\rangle$*, Theorem 4 implies that solutions of Problem (10) are same for* $T = T^{WD}, T^I$*, and* T^{JD}*, i.e.,*

$$s^*_{T^{WD}} = s^*_{T^{JD}} = s^*_{T^I}.$$

3.2 Implications of zero weights on input-efficiency improvements.

For a more focused analysis that addresses environmental problems created by the thermal power sector, our empirical work will be restricted to cases where zero weights are assigned to non-emission causing inputs while computing and analysing efficiency indexes. Theorem 7 below studies the consequences of zero weighting restrictions. In particular, the optimum of Problem

[22] The formulation below is an extension of Färe and Lovell (1978) and Färe et al (1985) to include the bad output dimension. See also Murty and Russell (2017).

[23] Moreover, if v is itself a strictly efficient point of T, then $s^*_T = 0_{n+2}$.

(10) recommends no efficiency improvements in usage of the non-emission causing inputs, when zero weights are assigned to these inputs under all the technology-modelling approaches studied in this paper.

For any technology set $T \subset \mathbf{R}_{+}^{n+2}$, define production function: $\Psi_T : \mathbf{R}_{+}^{n+1} \longrightarrow \mathbf{R}_{+}$ with image[24]

$$\Psi_T(x_o, x_z, z) := \max\{y \geq 0 \mid \langle x_o, x_z, y, z\rangle \in T\}.$$

Theorem 7 *Let T satisfy free input disposability for some input $i \in \{1, \ldots, n\}$. Let the initial production vector be $v = \langle x_o, x_z, y, z\rangle \gg 0_{n+2}$.*

(i) There exists a solution $s^* = \langle \delta_o^*, \delta_z^*, \theta^*, \gamma^*\rangle$ of Problem (10) with $\delta_i^* = 0$ whenever $w_i^x = 0$.[25]

(ii) If Ψ_T is strictly increasing in input i and $s^* = \langle \delta_o^*, \delta_z^*, \theta^*, \gamma^*\rangle$ is a solution of Problem (10) then $\delta_i^* = 0$ whenever $w_i^x = 0$.

The corollary below follows as, under the weak disposability-based and input approaches to modelling technologies, all inputs are freely disposable; while under the by-production and the joint production-based approaches, only the non-emission causing inputs are freely disposable.

Corollary 8 *The conclusions of Theorem 7 are true*

(i) for all inputs when $T \in \{T^{WD}, T^I\}$.

(ii) for all non-emission causing inputs ($i = 1, \ldots, n_o$) when $T \in \{T^{BP}, T^{JD}\}$.

3.3 The weighted output-based index of efficiency improvements.

The following problem defines the output-based weighted index of efficiency improvements $\mathcal{I}_T^O : T \times \Delta \longrightarrow \mathbf{R}_{+}$ for technology $T \in \{T^{BP}, T^{WD}, T^I\}$:

$$\begin{aligned} \mathcal{I}_T^O(x_o, x_z, y, z; w) \quad = \quad & \max_{\langle \delta_o, \delta_z, \theta, \gamma\rangle} \sum_{i=1}^{n_o} w_{o_i}^x \delta_{o_i} + \sum_{i=1}^{n_z} w_{z_i}^x \delta_{z_i} + w^y\theta + w^z\gamma \\ & \text{subject to} \\ & \big\langle x_o - (\delta_o \otimes x_o), \quad x_z - (\delta_z \otimes x_z), \quad y + \theta y, \quad z - \gamma z\big\rangle \in T, \\ & \delta_o = 0_{n_o}, \quad \delta_z = 0_{n_z}, \quad \gamma \in [0,1], \quad \theta \geq 0. \end{aligned} \tag{11}$$

[24]We assume that for any $\langle x_o, x_z, z\rangle \in \mathbf{R}_{+}^{n+1}$, the output-possibility set $P(x_o, x_z, z) = \{y \geq 0 \mid \langle x_o, x_z, y, z\rangle \in T\}$ is closed and bounded.

[25]That is, there may be multiple solutions of Problem (10), in which case, there is always a solution with $\delta_i^* = 0$.

Note that Problem (11) is just Problem (10) with an extra restriction that there is no change in all inputs, *i.e.*, $\delta_o = 0_{n_o}$ and $\delta_z = 0_{n_z}$.

The following remark states that the output-based weighted index of efficiency improvements is the same for the input and weak disposability-based technology modelling approaches.

Remark 9 *Suppose θ^O_κ and γ^O_κ solve (11) for technology T^κ. Then the point $\left\langle y + y\theta^O_\kappa, z - z\gamma^O_\kappa \right\rangle$ lies on the strictly efficient frontier of the output-possibility set $P^\kappa(x)$. It follows from conclusion (ii) of Theorem 4 that optimal output-based efficiency improvements are the same for the input and weak disposability-based approaches, i.e., $\left\langle \theta^O_{WD}, \gamma^O_{WD} \right\rangle = \left\langle \theta^O_I, \gamma^O_I \right\rangle$.*

3.4 Efficiency improvements in outputs attributable solely to efficiency improvements in input usage.

Suppose $\langle \theta^O, \gamma^O \rangle$ is a solution of Problem (11) for technology T. Since all inputs are held fixed in Problem (11), point $v^O := \left\langle x_o,\ x_z,\ y + \theta^O y,\ z - \gamma^O z \right\rangle$ lies on the weakly efficient frontier of technology T. On the other hand, as discussed earlier, if $s^* = \langle \delta^*_o, \delta^*_z, \theta^*, \gamma^* \rangle$ solves the weighted graph inefficiency Problem (10) when all inputs are allowed to vary, then $v^* = \left\langle x_o - (\delta^*_o \otimes x_o),\ \ x_z - (\delta^*_z \otimes x_z),\ \ y + \theta^* y,\ \ z - \gamma^* z \right\rangle$ is a strictly efficient point of T. Hence, the movement from the initial production vector $v = \langle x_o, x_z, y, z \rangle$ to v^* can be decomposed into

- a movement from v to v^O that puts the producing unit on the weakly efficient frontier of T with no change in input usage and
- a movement along the weakly efficient frontier of T from v^O to the strictly efficient point v^* following the optimal proportional reduction in input usage, δ^*.

Define

$$\gamma^x := \gamma^* - \gamma^O \quad \text{and} \quad \theta^x := \theta^O - \theta^*. \tag{12}$$

Then γ^x and θ^x can be interpreted as measuring, respectively, the optimal efficiency improvement (or the proportional *reduction*) in emission generation and the optimal efficiency improvement (or proportional *increase*) in production of the good output attributable solely to the vector of proportional reductions in usage of inputs, δ^*.

It is intuitive that the reduction in usage of inputs, including emission-causing inputs, should, ceteris paribus, also imply reductions in both the minimum level of emission that can be generated and the maximum amount of good output that can be produced. Hence, our intuition would suggest that (a) the *graph* measure of proportional *reduction* in emission that allows reductions in input usage should be no smaller that than the *output-based* measure that holds inputs fixed (*i.e.*, $\gamma^x \geq 0$) and (b) that the output-based measure of proportional *increase* in the intended output should be no smaller than the corresponding graph measure (*i.e.*, $\theta^x \geq 0$).

4 Efficiency improvements in outputs as functions of efficiency improvements in input usage.

Optimal efficiency improvements in production of the goods and bads can be shown to be functions of efficiency improvements in inputs. To see this, we first denote the sets $[0,1]^{n_z}$ and $[0,1]^n$ by $\mathcal{Q}_z$ and $\mathcal{Q}$, respectively, and define some additional technological constructs: For $\kappa = BP,\ WD,\ I,\ JD$, define the input requirement set

$$L^{\kappa}(y,z) = \left\{ \langle x'_o, x'_z \rangle \in \mathbf{R}^n_+ \;\middle|\; \langle x'_o, x'_z, y, z \rangle \in T^{\kappa} \right\}.$$

This is the set of all input bundles that can produce y and z levels of the good and the bad outputs, respectively, under technology T^{κ}. We now define the set

$$L^{\kappa}_{\delta} \;=\; \left\{ \langle \delta_o, \delta_z \rangle \in \mathcal{Q} \;\middle|\; \langle x_o - (\delta_o \otimes x_o)\,,\ x_z - (\delta_z \otimes x_z) \rangle \in L^{\kappa}(y,z) \right\} \tag{13}$$

as the set of proportional changes in inputs taking values in the set $\mathcal{Q}$ such that the corresponding changed levels of input can continue producing good and bad output levels y and z under technology T^{κ}. Note that, since $v = \langle x_o, x_z, y, z \rangle \in T^{\kappa}$, we have $\langle x_o, x_z \rangle \in L^{\kappa}(y,z)$. Hence, 0_n is in L^{κ}_{δ}.

At a solution $s^* = \langle \delta^*_o, \delta^*_z, \theta^*, \gamma^* \rangle$ of Problem (10) for technology T^{κ}, $\langle \delta^*_o, \delta^*_z \rangle$ must lie in L^{κ}_{δ}.[26] This implies that Problem (10) can be decomposed into

$$\begin{aligned} \mathcal{I}^G_{T^{\kappa}}(x_o, x_z, y, z; w) \;&=\; \max_{\delta_o, \delta_z} \sum_{i=1}^{n_o} w^x_{o_i} \delta_{o_i} + \sum_{i=1}^{n_z} w^x_{z_i} \delta_{z_i} + w^y \Theta^{\kappa}(\delta_o, \delta_z; w^y, w^z) + w^z \Gamma^{\kappa}(\delta_o, \delta_z; w^y, w^z) \\ & \quad \text{subject to} \\ & \quad \langle \delta_o, \delta_z \rangle \;\in\; L^{\kappa}_{\delta}, \end{aligned} \tag{14}$$

where, the functions $\Theta^{\kappa} : \mathcal{Q} \times [0,1]^2 \longrightarrow \mathbf{R}_+$ and $\Gamma^{\kappa} : \mathcal{Q} \times [0,1]^2 \longrightarrow [0,1]$ with images

[26] For if this were not so, then $\langle \delta^*_o, \delta^*_z \rangle \in \mathcal{Q} \setminus L^{\kappa}_{\delta}$. Hence, $\langle x_o - (\delta^*_o \otimes x_o)\,,\ x_z - (\delta^*_z \otimes x_z) \rangle$ cannot produce intended output levels y or greater. Hence, if $\left\langle x_o - (\delta^*_o \otimes x_o)\,,\ \ x_z - (\delta^*_z \otimes x_z)\,,\ \ y + \theta^* y,\ \ z - \gamma^* z \right\rangle \in T_{\kappa}$, then $\theta^* < 0$. Since $\theta^* < 0$ does not satisfy the constraints in Problem (10), s^* could not be a solution of Problem (10), which is a contradiction.

$\theta = \Theta^{\kappa}(\delta_o, \delta_z; w^y, w^z)$ and $\gamma = \Gamma^{\kappa}(\delta_o, \delta_z; w^y, w^z)$, respectively, solve the problem

$$\max_{\langle\theta,\gamma\rangle\in\mathbf{R}^2}\{w^y\theta + w^z\gamma \mid \left\langle x_o - (\delta_o \otimes x_o),\ x_z - (\delta_z \otimes x_z),\ y + \theta y,\ z - \gamma z\right\rangle \in T^{\kappa}\}. \tag{15}$$

Problem (15) identifies optimal efficiency improvements in intended output production and emission generation for every vector of proportional reductions in inputs. Problem (15) hence yields optimal efficiency improvements in production of the good and bad as functions of efficiency improvements in inputs. Problem (14) then solves for the optimal proportional reductions in inputs.

Remark 10 *If $s^*_\kappa = \langle\delta^*_{o\kappa}, \delta^*_{z\kappa}, \theta^*_\kappa, \gamma^*_\kappa\rangle$ solves Problem (10) for technology T^κ, then $\langle\delta^*_{o\kappa}, \delta^*_{z\kappa}\rangle$ solves Problem (14) with*

$$\theta^*_\kappa = \Theta^{\kappa}(\delta^*_{o\kappa}, \delta^*_{z\kappa}; w^y, w^z) \quad \text{and} \quad \gamma^*_\kappa = \Gamma^{\kappa}(\delta^*_{o\kappa}\delta^*_{z\kappa}; w^y, w^z). \tag{16}$$

The following remark states that the output-based weighted index of efficiency improvements is a special case of Problem (15) when $\langle\delta_o, \delta_z\rangle$ is chosen to be the zero vector $\langle 0_{n_o}, 0_{n_z}\rangle$.

Remark 11 *Suppose θ^O_κ and γ^O_κ solve (11) for technology T^κ. Then*

$$\Theta^{\kappa}(0_{n_o}, 0_{n_z}; w^y, w^z) = \theta^O_\kappa \quad \text{and} \quad \Gamma^{\kappa}(0_{n_o}, 0_{n_z}; w^y, w^z) = \gamma^O_\kappa, \quad \forall\ \kappa \in \{WD, JD, I, BP\}$$

so that the weighted output-based index of efficiency improvements is

$$\mathfrak{I}^O_{T^\kappa}(x_o, x_z, y, z; w) \ = \ w^y\theta^O_\kappa + w^z\gamma^O_\kappa.$$

4.1 The special case of by-production technologies.

The independence conditions (ii) and (i') in Definition 1 of a BPT imply that functions $\Theta^{BP}(\delta_o, \delta_z; w^y, w^z)$ and $\Gamma^{BP}(\delta_o, \delta_z; w^y, w^z)$ defined in (15) solve the separate problems

$$\begin{aligned}\Theta(\delta_o, \delta_z) := \Theta^{BP}(\delta_o, \delta_z; w^y, w^z) \ &= \ \max_{\theta\in\mathbf{R}}\{\theta \mid \left\langle x_o - (\delta_o \otimes x_o), x_z - (\delta_z \otimes x_z), y + \theta y, z\right\rangle \\ &\qquad \in T_1\} \qquad (17)\\ \Gamma(\delta_z) := \Gamma^{BP}(\delta_o, \delta_z; w^y, w^z) \ &= \ \max_{\gamma\in\mathbf{R}}\{\gamma \mid \left\langle x_o, x_z - (\delta_z \otimes x_z), y,\ z - \gamma z\right\rangle \\ &\qquad \in CDH(T_2)\}. \qquad (18)\end{aligned}$$

Function Θ^{BP} gives the maximum efficiency improvement in production of the intended output that is feasible under sub-technology T_1 for every vector of efficiency improvements in inputs

$\delta \in \Omega$ starting from the initial input vector x. Thus, this function is independent of the weights w^y and w^z and hence, in Problem 17, we redefine it as function $\Theta : \Omega \longrightarrow \mathbf{R}_+$, whose values depend only on $\langle \delta_o, \delta_z \rangle$. Similarly, function Γ^{BP}, which gives the maximum efficiency improvement in the generation of emission that is feasible under sub-technology $CDH(T_2)$ for every given vector of efficiency improvements in emission-causing inputs $\delta_z \in \Omega_z$, is redefined in Problem 18 as function $\Gamma : \Omega_z \longrightarrow [0, 1]$ that is independent of the weights.[27]

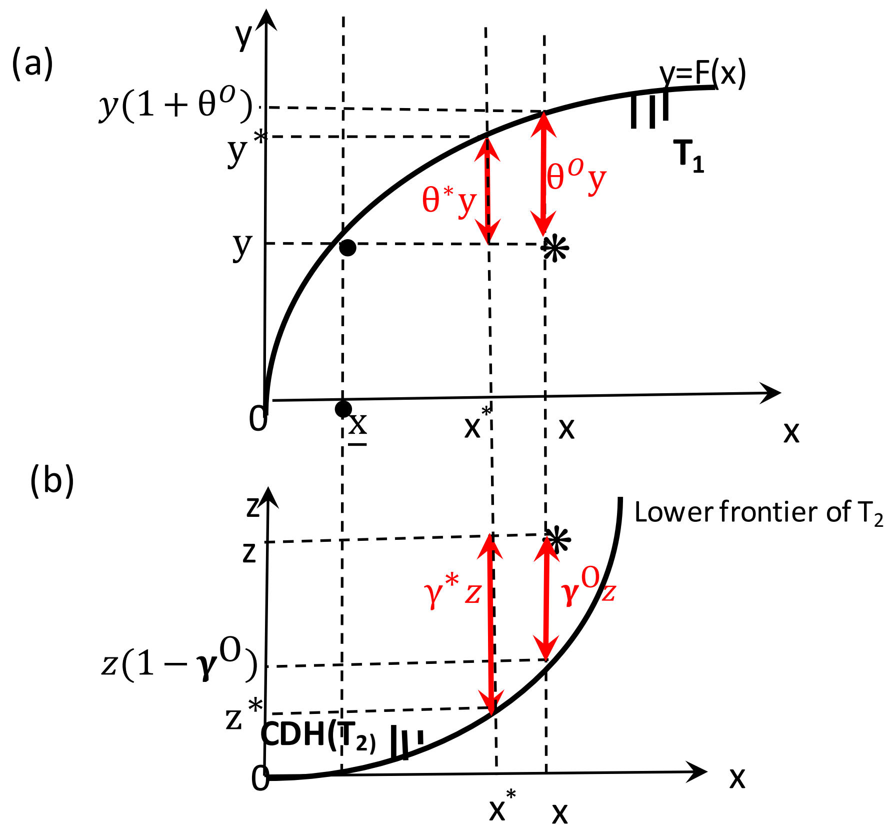

Figure 2

Figure 2 explains the intuition behind functions Θ and Γ. It assumes that there is a single input that is emission-causing, *i.e.*, $n = n_z = 1$. Starting from x, if the input vector is reduced by a proportion equal to $\delta^* \in \Omega$, then the new vector of inputs is $x^* = x - \delta^* x < x$. Panel (a) of Figure 2 shows that the proportional reduction δ^* in the input vector implies that the maximum possible efficiency improvement in intended output production under sub-technology T_1 is $\Theta(\delta^*) = \theta^*$, and the new intended output level is $y^* = y + \theta^* y > y$. At the same time, panel (b) of Figure 2 shows that the proportional reduction δ^* in the input vector implies that the maximum possible efficiency improvement in the emission generation under sub-technology T_2 is $\Gamma(\delta^*) = \gamma^*$, and the new level of emission is $z^* = z - \gamma^* z < z$.

[27]Note, because of condition (i') in Definition 1 of a BPT, function Γ^{BP} is independent of the proportional changes in non-emission causing inputs. Hence, values of function Γ depend only of δ_z.

In Figure 2, suppose $s^* = \langle \delta^*, \theta^*, \gamma^* \rangle$ solves the graph efficiency improvement Problem (10). Then the amount of input reduction is $\Delta x = x - x^*$. The total reduction in emission, $\gamma^* z$, can be decomposed into (i) $\gamma^O z$, the maximum reduction in emission possible when there is no change in input usage and (ii) $\left(\gamma^* - \gamma^O\right) z$, the maximum reduction in emission attributable to reduction in input usage Δx. The loss in potential for efficiency improvement in good output production due to reduction in input Δx is given by $\left(\theta^O - \theta^*\right) y$, which is equivalent to the difference between (i) $\theta^O y$, the maximum increase in good output that is possible with no change in input usage and (ii) $\theta^* y$, the maximum increase in good output that is possible when there is a reduction in the input usage by an amount Δx.

4.2 Exploring monotonicity properties of functions Θ^κ and Γ^κ.

Theorem 12, whose proof can be found in the appendix, describes the monotonicity and curvature properties of functions Θ^κ and Γ^κ for $\kappa \in \{WD, JD, I, BP\}$.

Theorem 12 *The following are true:*

(i) $w^y \Theta^{WD}(\delta_o, \delta_z; w^y, w^z) + w^z \Gamma^{WD}(\delta_o, \delta_z; w^y, w^z)$ *is a non-increasing function of* $\langle \delta_o, \delta_z \rangle \in \Omega$.

(ii) Suppose both $\langle \hat{\delta}_o, \hat{\delta}_z \rangle$ *and* $\langle \bar{\delta}_o, \bar{\delta}_z \rangle$ *are in* Ω *such that* $\langle \hat{\delta}_o, \hat{\delta}_z \rangle \leq \langle \bar{\delta}_o, \bar{\delta}_z \rangle$.
Let $\hat{\theta}^{JD} = \Theta^{JD}\left(\hat{\delta}_o, \hat{\delta}_z; w^y, w^z\right)$, $\bar{\theta}^{JD} = \Theta^{JD}\left(\bar{\delta}_o, \bar{\delta}_z; w^y, w^z\right)$, $\hat{\gamma}^{JD} = \Gamma^{JD}\left(\hat{\delta}_o, \hat{\delta}_z; w^y, w^z\right)$, *and* $\bar{\gamma}^{JD} = \Gamma^{JD}\left(\bar{\delta}_o, \bar{\delta}_z; w^y, w^z\right)$. *Suppose* $\bar{v}$ *and* $\hat{v}$ *are strictly efficient points of* T^{JD}, *where*
$\bar{v} = \left\langle x_o - \left(\bar{\delta}_o \otimes x_o\right),\ x_z - \left(\bar{\delta}_z \otimes x_z\right),\ y + \bar{\theta}^{JD} y,\ z - \bar{\gamma}^{JD} z \right\rangle$ *and*
$\hat{v} = \left\langle x_o - \left(\hat{\delta}_o \otimes x_o\right),\ x_z - \left(\hat{\delta}_z \otimes x_z\right),\ y + \hat{\theta}^{JD} y,\ z - \hat{\gamma}^{JD} z \right\rangle$. *Then*

$$w^y \hat{\theta}^{JD} + w^z \hat{\gamma}^{JD} \geq w^y \bar{\theta}^{JD} + w^z \bar{\gamma}^{JD}$$

(iii) $\Theta^{BP} \equiv \Theta$ *is a non-increasing function defined on domain* Ω *and* $\Gamma^{BP} \equiv \Gamma$ *is a non-decreasing function defined on domain* Ω_z. *Both functions are concave.*

While the by-production approach implies that both functions Θ and Γ are individually monotonic (albeit in different directions), the weak disposability-based approach implies non-positive monotonicity of only the weighted average of functions Θ^{WD} and Γ^{WD} and the joint disposability-based approach implies non-positive monotonicity of the weighted average of functions Θ^{JD} and Γ^{JD} only along the strictly efficient frontier of the technology, *i.e.,* only to the extent efficiency improvements lead to the strictly efficient frontier of the technology.

Thus, the above theorem highlights the trade-off (negative correlation) between efficiency improvement in the good and bad output directions under the by-production approach when inefficiency is removed in the direction of the emission-causing inputs: Ceteris-paribus, under this approach, when there is a proportionate reduction in the usage of inputs, the maximum extent of efficiency improvement in production of the good output cannot increase (it can go down), while the maximum extent of efficiency improvement in generation of the bad output cannot decrease (it can go up). Figure 2 shows that when δ increases from zero to δ^*, the value of function Θ decreases from θ^O to θ^* and the value of function Γ increases from γ^O to γ^*.

In contrast, under the weak and joint disposability-based approaches, conclusions (i) and (iii) of Theorem 12 are consistent with the maximum extent of efficiency improvement in emission generation going down and the maximum extent of efficiency improvement in the good output production going up when inefficiencies are removed in the input direction. That is, it is possible that $\hat{\delta} < \bar{\delta}$ implies $\hat{\theta} < \bar{\theta}$ and $\hat{\gamma} > \bar{\gamma}$.

5 Data and implications of emission-data generating process.

5.1 Data description.

This study uses data on 47 coal-fired thermal power plants in India for the year 2014. The data was collected from the annual publication of the Central Electricity Authority (CEA) of India (2013-14, 2014-15). The plants studied are run by 16 major power generating companies operating in various states of India. For greater details on the dataset, the reader is referred to Murty and Nagpal (2018).

The intended output of the power plants is net electricity, which is measured in gigawatt hours (GWh). Since data on capital and labour employed in these power plants is not available, plant capacity, measured in megawatt (MW), is used as a proxy for capital,[28] while labour is not included as an input.[29]

Since data on oil consumption measured in heat units is sporadic and incomplete, aggregate heat from coal and oil consumption by the coal-based thermal power plants, denoted by x_z and measured in millions of kilocalories (mill of Kcal), is taken as the only emission-causing input (fossil fuel input).[30] This can be obtained by multiplying the data on station heat rate (SHR)

[28] A similar approach is also taken in other recent works on Indian thermal power plants (see *e.g.,* Sahoo et al (2017) and Behera et al (2010)).

[29] It has been argued that the contribution of labour cost to total operating costs of these power plants is very small (see *e.g.,* Kumar et al (2015)).

[30] Oil is the secondary fuel in coal-based thermal power plants. It is employed primarily to cover the start-up fuel requirements and for flame stabilisation, and does not contribute significantly to electricity generation in these power plants.

with that on gross electricity generation.[31]

The model also employs CEA data on plant operating availability as a managerial input.[32] It is the percentage of total capacity (measured in MWh) that is available to the plant for electricity generation after subtracting out the percentage lost due to forced outage and planned maintenance.

Data on emissions is not generated by direct observation. Rather, the following linear deterministic formula is used by CEA to compute the emission level (measured in metric tons (Mtons) from fossil-fuel consumption by plant u, converted into heat units:

$$\begin{aligned} z^u &= [(\text{heat from coal})^u \times EF_C \times Oxid_C] + [(\text{heat from oil})^u \times EF_O \times Oxid_O] \\ &= \left[x^u_{z_C} \times EF_C \times Oxid_C\right] \quad + \quad \left[x^u_{z_O} \times EF_O \times Oxid_O\right], \end{aligned} \tag{19}$$

where for fuel type $i =$ coal (C) or oil (O), x_{z_i} denotes the heat from fossil fuel i and EF_i and $Oxid_i$ are the emission and oxidation factors of the fossil-fuel i, respectively. EF_i is the amount of emission per unit of heat generated by burning fossil-fuel of type i. These are assumed to be constant across all plants with EF_C and $Oxid_C$ taking values 92.5 grams per megajoules (g/MJ) and .98, respectively; while $Oxid_O$ takes a value equal to one and EF_O takes a value 71.9 g/MJ for all plants.[33]

5.2 Almost perfect correlation between data on emission and aggregate heat input.

Since comprehensive data on heat from oil is not available, in our study, data on emission generation by coal-fired power plants is related to aggregate heat from oil and coal, which is our only emission-causing input. Since the emission and oxidation factors of oil are lower than that of coal, (19) implies that, among plants employing the same amount of aggregate heat input, $x_z = x_{z_C} + x_{z_O}$, those having a higher share of oil viz-a-viz coal will have lower levels of emission.

Murty and Nagpal (2018) observe that the correlation between data on emission generation and the aggregate heat employed by power plants in our data set is 0.9993 in year 2014.[34] This

[31]The station heat rate is defined as the amount of heat required by a power plant to produce one unit of electricity. It is computed by CEA as $SHR = (SCC \times GCV_c) + (SOC \times GCV_o)$, where GCV_c and GCV_o are the gross calorific values of coal and oil respectively, while SCC and SOC are the specific coal and oil consumption respectively. SCC measures the physical units of coal required to generate one unit of electricity and GCV_c is the heat content of a physical unit of coal. SOC and GCV_o are similarly defined.

[32]Many recent works on thermal power plants such as Sahoo et al (2017) and references there in include managerial inputs.

[33]See Appendix B of the user guide of CO_2 baseline database for the Indian power sector, published by CEA.

[34]See Table 2 of Murty and Nagpal (2018).

near perfect correlation has been attributed to the tiny share of oil in the total heat employed by the power plants.[35] Thus, heat from oil contributes insignificantly to emission generation by coal-based power plants and most of the emission generated can be attributed to usage of coal. Hence, it follows from (19) that the relation between data on aggregate heat and emission generation will be nearly linear (proportional), with the reference constant of proportionality being given by the emission factor of coal, $EF_C \times Oxid_C$.

6 A comparison of efficiency improvements under different technology-modelling approaches: The case of Indian coal-based thermal power sector.

The CEA dataset on Indian coal-based thermal power sector offers a unique test for comparing and evaluating the performance of the various approaches to modelling emission-generating technologies studied in Section 2. Since the relation between CEA data on heat input and emission generation is nearly linear, the predicted changes in the amounts of emission generated due to changes in the amounts of heat input employed by power plants will have to bear an almost proportional relation to the latter changes, with the reference constant of proportionality being $EF_C \times Oxid_C$. In other words, the reduction in emission attributed solely to reduction in the heat input *per unit* reduction in the heat input has to be close to the factor $EF_C \times Oxid_C$.

In the empirical analysis below, we study to what extent this is true of the optimal efficiency improvements derived from solving Problem (10) and (11) under the by-production (BP), weak disposability (WD), and joint disposability (JD)-based modelling approaches, as both the graph-based and output-based optimal efficiency improvements for WD and input approach-based models are identical.[36] As an illustrative example, an equal weight of one-third is assigned to efficiency improvements in usage of the heat input, production of electricity, and generation of CO_2, while solving Problems (10) and (11).[37]

[35]See Table 4 in Murty and Nagpal (2018), which provides evidence based on the limited available data on oil.

[36]See Theorem 4 and Remark 11.

[37]This implies that zero weights are assigned to efficiency improvements in the usage of the two non-emission causing inputs. The implications of this choice of weights is shown in Theorem 7 and its corollary.

6.1 Comparison of optimal efficiency improvements under different technological specifications.

Tables 1 and 2 show some basic descriptive statistics of optimal efficiency improvements and the graph and output-based efficiency indexes computed by solving Problems (10) and (11).[38] As seen in Table 1, the descriptive statistics of graph efficiency and optimal efficiency improvements (θ^*, γ^*, and δ_z^*) obtained from solving Problem (10) for the WD and JD-based models are identical. This is in confirmation with our theoretical observation summarised in Remark 6 that technologies based on both these approaches have identical strictly efficient frontiers. However, this table shows considerable variations across all three models with respect to several output-based measures computed by solving Problem (11).

Table 1

	BP				WD				JD			
Graph	**EI**	**θ***	**γ***	**δ_z***	**EI**	**θ***	**γ***	**δ_z***	**EI**	**θ***	**γ***	**δ_z***
max	0.968	2.712	0.579	0.553	1	2.624	0.499	0.553	1	2.624	0.499	0.553
min	0.516	0	0.094	0	0.529	0	0	0	0.529	0	0	0
mean	0.858	0.097	0.264	0.180	0.885	0.075	0.172	0.191	0.885	0.075	0.172	0.191
median	0.857	0	0.261	0.165	0.882	0	0.161	0.180	0.882	0	0.161	0.180
std dev	0.090	0.429	0.131	0.151	0.095	0.399	0.139	0.154	0.095	0.399	0.139	0.154
Output-based	**EI**	**θ^O**	**γ^O**	**b_y**	**EI**	**θ^O**	**γ^O**		**EI**	**θ^O**	**γ^O**	
max	0.968	2.712	0.181	1	1	2.624	0.291		1	2.309	0.077	
min	0.516	0	0	0.2694	0.533	0	0		0.565	0	0	
mean	0.899	0.276	0.100	0.883	0.927	0.183	0.089		0.949	0.184	0.004	
median	0.916	0.149	0.102	0.870	0.938	0	0.048		0.974	0.081	0	
std dev	0.083	0.442	0.027	0.162	0.083	0.438	0.096		0.074	0.356	0.015	

EI: efficiency Index

Although the rank correlation for graph efficiency index between the WD (and hence JD) and BP-based models is high (0.995) and Table 1 shows that the optimal efficiency improvement in the usage of the heat input are similar across all models (mean value of δ_z^* lies between 18% and 19% across all three models), the mean value of the optimal proportional reduction in emission attributable solely to efficiency improvement in usage of heat input ($\gamma^* - \gamma^O$) is higher for the JD and BP models as opposed to the WD model. In fact, the mean values of $\gamma^* - \gamma^O$ computed for the JD and BP models are very close (16.8% and 16.4%, respectively) and more than double the value for the WD model. Thus, in the context of our dataset, *the WD-based model does not effectively translate high proportional reductions in the heat input into high proportional reductions in emission generation.*

There are also notable differences in the composition of the difference $\gamma^* - \gamma^O$ across JD and BP models. Both the graph and output-based measures γ^* and γ^O are higher for BP model as

[38] The graph efficiency index is computed as $\frac{1}{1+\mathfrak{I}_T^G(x_o,x_z,y,z;w)}$, where $\mathfrak{I}_T^G(x_o,x_z,y,z;w)$ is the weighted graph index of efficiency improvements computed by Problem (10). Similarly, too we can compute the output-based efficiency index from the weighted output-based index of efficiency improvements computed by Problem (11).

compared to the JD model.

6.2 Verifying monotonicity properties of optimal efficiency improvements in production of goods and bads.

Since $\delta_z^* \geq \delta_z^O = 0$, for the WD model, Theorem 12 would imply

$$w^y \left[\theta^O - \theta^*\right] \geq w^z \left[\gamma^* - \gamma^O\right]. \tag{20}$$

Since w^y is assumed to be equal to w^z in our empirical analysis, Table 2 demonstrates that the above holds for our dataset with both $\theta^O - \theta^*$ and $\gamma^* - \gamma^O$ taking positive values. For the BP case Theorem 12 would imply $\theta^* \leq \theta^O$ and $\gamma^* \geq \gamma^O$, both of which, as seen in Table 2, are validated by our dataset. Hence, in both WD and BP cases, our dataset implies that as efficiency improvement in usage of heat input increases, the optimal efficiency improvement in emission generation increases while the optimal efficiency improvement in electricity generation falls. In the BP case, this especially confirms to our real-world based intuition: as usage of emission-causing heat input decreases, both the minimum amount of emission and the maximum amount of electricity that can be produced fall (See Figure 2).

Table 2

	γ^*-γ^O			θ^O-θ^*		
	WD	JD	BP	WD	JD	BP
max	0.499	0.499	0.520	0.795	0.500	0.744
min	0	0	0	0	-0.316	0
mean	0.083	0.168	0.164	0.108	0.108	0.179
median	0	0.144	0.150	0	0.073	0.125
std dev	0.149	0.138	0.141	0.205	0.145	0.192

Table 3

$\Delta z/\Delta x_z$	BP	WD	JD	z/x_z
min	0.339	0	0	0.339
max	0.339	0.372	0.53157	0.414
mean	0.339	0.090	0.3256179	0.377
median	0.339	0	0.342114	0.378
std dev	0	0.142	0.072039	0.012
Correl $\Delta z,\Delta x_z$	1	0.620	0.989	0.9993 (Correl z,x_z)

Table 2 also shows that, for the JD model, there exist plants in our dataset for which $\theta^O - \theta^*$ is negative, while $\gamma^* - \gamma^O$ is positive.[39]

Table 2 also shows that, for the JD model, there exist plants in our dataset for which $\theta^O - \theta^*$ is negative, while $\gamma^* - \gamma^O$ is positive.[40] This implies that, holding all other inputs fixed, as usage of the heat input falls, the optimal amount of electricity production recommended by the JD model increases, while the optimal amount of emission generation recommended by it falls. While the latter is intuitive, the former may appear counterintuitive as reduction in usage of heat input, holding all other inputs fixed, can normally be expected to imply a fall in the level of maximum electricity that can be produced.

[39] This is not in conflict with (20) or conclusions of Theorem 12. Rather, it implies that, for plants for which $\theta^O - \theta^*$ is negative and $\gamma^* - \gamma^O$ is positive, the efficiency improvement vector $\langle 0_n, \theta^O, \delta^O \rangle$ does not lead to a point on the strictly efficient frontier of technology T^{JD}.

[40] This is not in conflict with (20) or conclusions of Theorem 12. Rather, it implies that, for plants for which $\theta^O - \theta^*$ is negative and $\gamma^* - \gamma^O$ is positive, the efficiency improvement vector $\langle 0_n, \theta^O, \delta^O \rangle$ does not lead to a point on the strictly efficient frontier of technology T^{JD}.

6.3 Are recommended changes in emission levels proportional to recommended reductions in input usage?

Table 3 gives the descriptive statistics of emission reductions attributable solely to reduction in usage of heat input *per unit* reduction in usage of the heat input ($\Delta z/\Delta x_z$) across the three models. Here, $\Delta z = \left(\gamma^* - \gamma^O\right) \times z$, while $\Delta x_z = \delta_z \times x_z$, where z and x_z are the starting amounts of emission and the heat input. As discussed above, given the way emission data is generated by CEA, one expects the ratio $\Delta z/\Delta x_z$ to be nearly constant across all production units. Table 3 reveals that for the BP model this ratio is exactly a constant equal to 0.339, which is not far from the reference factor $EF_C \times Oxid_C = 0.3795$ Mtons/mill Kcal.[41] As seen in the table, there is a considerable variation in the range of values $\Delta z/\Delta x_z$ takes across power plants in the WD and JD-based models. For example, for the JD model the values vary from a minimum of zero to a maximum of 0.532.

Table 4

Plants with $\delta_z^* = 0$	WD and JD				BP			
Name	δ_z^*	θ^*	γ^*	EI	δ_z^*	θ^*	γ^*	EI
BOKARO B	0	0	0	1	0	0	0.160	0.949
DAHANU	0	0	0	1	0	0	0.106	0.966
KORBA	0	0	0	1	0	0	0.102	0.967
VINDHAYACHAL	0	0	0	1	0	0	0.100	0.968
No. of plants with with $\delta_z^* > 0$ and $\gamma^*-\gamma^O=0$	WD	JD	BP					
	30	1	0					

Moreover Table 4 shows that, under the WD approach, the optimum of Problem (10) recommends positive proportional reductions in heat input usage (*i.e.,* $\delta_z > 0$) with no accompanying reduction in emission generation (*i.e.,* $\gamma^* - \gamma^O = 0$) for thirty plants (which form nearly 75% of plants in the dataset). Under the JD approach there is one plant with this feature. This is again counterintuitive, as the nature of CEA's data generating process for emission implies that changes in changes in the usage of the heat input should induce nearly proportional changes in the amount of emission generated. Under the BP approach, positive proportional reductions in heat input usage are always accompanied by positive proportional reductions in emission generation.

[41] In fact Murty and Nagpal (2018) show that 0.339 is the slope of the lower frontier of sub-technology T_2 in the emission-heat input space. This frontier is linear because of the constant returns to scale assumption. If non-increasing returns to scale is assumed, then too the slope of the efficient frontier of T_2 under the BP approach is expected to show very little variations.

6.4 Benchmark production vectors under alternate modelling approaches.

Table 4 also shows that while four plants have graph efficiency equal to one under the JD and WD approaches (*i.e.,* the initial production vectors of these plants lie on the strictly efficient frontiers of technologies T^{WD} and T^{JD}), there is no "peer group" of plants that operate on the strictly efficient frontier of T^{BP}.

Rather, under the BP approach, while each of these four plants operates efficiently along the heat-input and electricity dimensions (*i.e.,* $\delta_z^* = \theta^* = 0$), it is associated with inefficiency in the emission direction (*i.e.,* $\gamma^* > 0$). Since $\delta_z^* = 0$, this inefficiency is simply output-based, *i.e.,* $\gamma^O = \gamma^*$. In other words, at the initial production points of these plants, the emission generation levels are higher than the minimum possible that can be generated with their existing levels of usage of the heat input. Thus, while these plants operate on the efficient frontier of sub-technology T_1, they operate above the lower frontier of sub-technology T_2. Nevertheless, given the initial production vector $v = \langle x_o, x_z, y, z \rangle \in T^{BP}$ for each plant, the BP approach does identify production bundle $\langle x_o - (\delta_o^* \otimes x_o)\,,\ x_z - (\delta_z^* \otimes x_z)\,,\ y + \theta^* y,\ z - \gamma^* z \rangle$ as the efficient reference (benchmark) production bundle relative to which its graph technical efficiency is measured.

7 Tabulating all possible solutions of Problem (10) under the by-production approach.

We tabulate all possible solutions of Problem (10) under the BP approach. Given the separate monotonicity properties of functions Θ and Γ under this approach, the solutions reflect the trade-offs between efficiency improvements in good and bad outputs given efficiency improvements in input usage. In particular, under the BP approach, $\theta^O - \theta^*$ and $\gamma^* - \gamma^O$ are both non-negative whenever $\delta_z^* > 0$. This non-positive correlation may not be displayed by solutions obtained under other technological modelling approaches, as Theorem 12 establishes monotonicity of at most the weighted average of Θ^κ and Γ^κ for $\kappa = WD, JD$ and hence implies (20), which is consistent with a decrease in efficiency improvement in good output production ($\theta^O - \theta^* > 0$) being accompanied by a decrease in efficiency improvement in generation of the bad ($\gamma^* - \gamma^O < 0$) when there are efficiency improvements in usage of the inputs ($\delta_z^* > 0$).

7.1 Images of functions Θ and Γ.

We derive the images of functions Θ and Γ and then employ these to tabulate all possible solutions of Problem 10 for a BPT. In what follows we will assume that the initial production vector $v = \langle x_o, x_z, y, z\rangle$ lies in T^{BP}. To study the image of function Θ, we will first partition its domain Ω into four mutually exclusive and exhaustive sets. To this end, we first derive the following production function obtained by maximising the production of intended output under sub-technology T_1 for any given vector of input levels.[42]

$$F^{BP}(\bar{x}_o, \bar{x}_z) := \max\{\bar{y} \geq 0 \mid \langle \bar{x}_o, \bar{x}_z, \bar{y}, \bar{z}\rangle \in T_1\}.$$

Under the free disposability condition (i) in Definition 1, function F can be shown to be non-decreasing in all inputs.[43] The input requirement set of sub-technology T_1 corresponding to y level of the intended output is defined as

$$L^{BP}(y) = \left\{\langle \bar{x}_o, \bar{x}_z\rangle \in \mathbf{R}^n_+ \;\middle|\; y \leq F^{BP}(\bar{x}_o, \bar{x}_z)\right\}.$$

We now define the set

$$L^{BP}_\delta \quad = \quad \left\{\langle \delta_o, \delta_z\rangle \in \Omega \;\middle|\; y \leq F^{BP}\left(x_o - (\delta_o \otimes x_o)\,,\; x_z - (\delta_z \otimes x_z)\right)\right\} \tag{21}$$

For any vector of proportional changes in inputs $\langle \delta_o, \delta_z\rangle \in L^{BP}_\delta$, the input vector $\big\langle x_o - (\delta_o \otimes x_o)\,,\; x_z - (\delta_z \otimes x_z)\big\rangle$ lies in the input requirement set $L^{BP}(y)$.

The isoquant of sub-technology T_1 corresponding to y level of intended output is

$$I^{BP}(y) = \left\{\langle \bar{x}_o, \bar{x}_z\rangle \in \mathbf{R}^n_+ \mid F^{BP}(\bar{x}_o, \bar{x}_z) = y\right\}$$

The following subset of L^{BP}_δ is the set of all proportional changes in inputs taking values in set Ω that, starting from $\langle x_o, x_z\rangle$, lead to points in the isoquant of y :

$$I^{BP}_\delta = \left\{\langle \delta_o, \delta_z\rangle \in \Omega \;\middle|\; F^{BP}\left(x_o - (\delta_o \otimes x_o)\,,\; x_z - (\delta_z \otimes x_z)\right) = y\right\}.$$

Note that the set I^{BP}_δ may or may not include 0_n. If I^{BP}_δ includes 0_n then $F^{BP}(x_o, x_z) = y$ and the production vector $v = \langle x_o, x_z, y, z\rangle$ is a weakly efficient point of sub-technology T_1.

It follows from the above definitions that set L^{BP}_δ can be re-written as the following union of disjoint sets: $L^{BP}_\delta = I^{BP}_\delta \setminus \{0_n\} \;\bigcup\; \{0_n\} \;\bigcup\; L^{BP}_\delta \setminus \left(I^{BP}_\delta \cup \{0_n\}\right)$.

[42] Since T_1 satisfies independence from the emission level, z is not shown as an argument of function F.

[43] See, for instance, Russell (1998) and Murty and Russell (2017).

From this it follows that set Ω can be written as the union of four mutually exclusive and exhaustive sets:

$$\Omega \;=\; \Omega \setminus L_{\delta}^{BP} \;\bigcup\; I_{\delta}^{BP} \setminus \{0_n\} \;\bigcup\; \{0_n\} \;\bigcup\; L_{\delta}^{BP} \setminus \left(I_{\delta}^{BP} \cup \{0_n\}\right). \tag{22}$$

Remark 13 *(17) and (18) imply that* $\theta^* = \Theta\left(\delta_o, \delta_z\right)$ *and* $\gamma^* = \Gamma\left(\delta_z\right)$ *if and only if* [44]

$$\begin{aligned} y(1+\theta^*) &= F^{BP}\left(x_o - \left(\delta_o \otimes x_o\right), x_z - \left(\delta_z \otimes x_z\right)\right) \\ z(1-\gamma^*) &= G\left(x_z - \left(\delta_z \otimes x_z\right);\; CDH(T_2)\right) \end{aligned} \tag{23}$$

The values that function Θ takes in the four subsets of its domain Ω and the image of function Γ are given by the following theorem:

Theorem 14 *The image of function* Θ *is given by*

$$\begin{aligned} \Theta\left(\delta_o, \delta_z\right) &\in [-1, 0) && \text{if} \quad \langle \delta_o, \delta_z \rangle \in \Omega \setminus L_{\delta}^{BP} \\ &= 0 && \text{if} \quad \langle \delta_o, \delta_z \rangle \in I_{\delta}^{BP} \setminus \{0_n\} \\ &\in \left(0, \theta^O\right] && \text{if} \quad \langle \delta_o, \delta_z \rangle \in L_{\delta}^{BP} \setminus \left(I_{\delta}^{BP} \cup \{0_n\}\right) \\ &= \theta^O && \text{if} \quad \langle \delta_o, \delta_z \rangle = 0_n. \end{aligned} \tag{24}$$

The image of function Γ *is given by*

$$\begin{aligned} \Gamma\left(\delta_z\right) &\in [\gamma^O, 1] && \text{if} \quad \delta_z \in \Omega_z \setminus \{0_{n_z}\} \\ &= \gamma^O && \text{if} \quad \delta_z = 0_{n_z} \end{aligned} \tag{25}$$

As Figures 2 and 3 assume only one input that is emission-causing, $\Omega = \Omega_z = [0, 1]$. Here, $L^{BP}(y) = [\underline{\mathrm{x}}, \infty)$ and $I^{BP} = \{\underline{\mathrm{x}}\}$. Thus, $L_{\delta}^{BP} = [0, \underline{\delta}]$ and $I_{\delta}^{BP} = \{\underline{\delta}\}$, where $\underline{\delta}$ solves $\underline{x} = x - \underline{\delta} x$. Given a proportional change in input δ^*, we assume that $x^* = x - \delta^* x$. In panel (a) of Figures 2 and Figure 3, $\underline{x} < x^* < x$. Hence, δ^* must lie in $L_{\delta}^{BP} \setminus (I_{\delta}^{BP} \cup \{0\})$. Suppose $\theta^* = \Theta(\delta^*)$. In particular, in Figure 2, $0 < \theta^* < \theta^O = \theta^O$. In contrast, in panel (a) of Figure 3, $\theta^* = \theta^O = \theta^O > 0$. In panel (b) of Figure 3, $x^* = \underline{x} < x$. Hence, δ^* must lie in $I_{\delta}^{BP} \setminus \{0\}$. Here $\theta^* = \theta^O = \theta^O = 0$.[45]

In panel (b) of Figure 2 and panels (b) and (d) of Figure 3, δ^* (which is exactly as defined above) is in $\Omega \setminus \{0\}$. While in Figure 2, $\gamma^* > \gamma^O$; in Figure 3, $\gamma^* = \gamma^O$. In particular, in panel (b) of Figure 3, $\gamma^O > 0$; while in panel (d) of Figure 3, $\gamma^O = 0$.

[44] Function G is defined in (2).

[45] In contrast, if in panel (a) of Figure 2, x^* was coincident with $\underline{x}$, then δ^* would also lie in $I_{\delta}^{BP} \setminus \{0\}$. But here, $0 = \theta^* < \theta^O = \theta^O$.

7.2 Tabulating all possible solutions of Problem (10) for a BPT.

If $s^* = \langle \delta_o^*, \delta_z^*, \theta^*, \gamma^* \rangle$ denotes a solution of problem (10) for initial production vector $v = \langle x_o, x_z, y, z \rangle \gg 0_{n+2}$ and technology $T = T^{BP}$, then Remark 10 applied to a BPT implies that $\theta^* = \Theta(\delta_o^*, \delta_z^*)$ and $\gamma^* = \Gamma(\delta_z^*)$. Employing this observation, Table 5 provides the set of all possible solutions of problem (10) given the initial production vector v and technology T^{BP}.

The columns of Table 5 cover the possible values γ^* can take given the image of function Γ defined in (25), while its rows cover the possible values θ^* can take given the image of function Θ defined in (24). For each combination of values of γ^* and θ^*, the table identifies the possible values vector $\langle \delta_o^*, \delta_z^* \rangle$ can take such that $\theta^* = \Theta(\delta_o^*, \delta_z^*)$ and $\gamma^* = \Gamma(\delta_z^*)$.

	$\gamma > \gamma^O \geq 0$ **(1)**	$\gamma = \gamma^O > 0$ **(2)**	$\gamma = \gamma^O = 0$ **(3)**
$0 < \theta < \theta^O$ **(1)**	$\delta \in L_\delta^{BP} \setminus \left(I_\delta^{BP} \cup \{0_n\}\right)$ $\delta_z > 0_{n_z}$	$\delta \in L_\delta^{BP} \setminus \left(I_\delta^{BP} \cup \{0_n\}\right)$ $\delta_z = 0_{n_z}$ or $\delta_z > 0_{n_z}{}^\dagger$	$\delta \in L_\delta^{BP} \setminus \left(I_\delta^{BP} \cup \{0_n\}\right)$ $\delta_z = 0_{n_z}{}^\spadesuit$ or $\delta_z > 0_{n_z}{}^\dagger$
$0 = \theta < \theta^O$ **(2)**	$\delta \in I_\delta^{BP} \setminus \{0\}$ $\delta_z > 0_{n_z}$	$\delta \in I_\delta^{BP} \setminus \{0\}$ $\delta_z = 0_{n_z}$ or $\delta_z > 0_{n_z}{}^\dagger$	$0_n < \delta \in I_\delta^{BP} \setminus \{0\}$ $\delta_z = 0_{n_z}{}^\spadesuit$ or $\delta_z > 0_{n_z}{}^\dagger$
$\theta = \theta^O > 0$ **(3)**	$\delta = 0_n$ or $\delta \in L_\delta^{BP} \setminus \left(I_\delta^{BP} \cup \{0_n\}\right)^*$ $\delta_z > 0_{n_z}$	$\delta = 0_n$ or $\delta \in L_\delta^{BP} \setminus \left(I_\delta^{BP} \cup \{0_n\}\right)^*$ $\delta_z = 0_{n_z}$ or $\delta_z > 0_{n_z}{}^\dagger$	$\delta = 0_n$ or $\delta \in L_\delta^{BP} \setminus \left(I_\delta^{BP} \cup \{0_n\}\right)^*$ $\delta_z = 0_{n_z}{}^\spadesuit$ or $\delta_z > 0_{n_z}{}^\dagger$
$\theta = \theta^O = 0$ **(4)**	$\delta = 0_n{}^\clubsuit$ or $\delta \in I_\delta^{BP} \setminus \{0_n\}^*$ $\delta_z > 0_{n_z}$	$\delta = 0_n{}^\clubsuit$ or $\delta \in I_\delta^{BP} \setminus \{0_n\}^*$ $\delta_z = 0_{n_z}$ or $\delta_z > 0_{n_z}{}^\dagger$	$\delta = 0_n{}^\clubsuit$ or $\delta \in I_\delta^{BP} \setminus \{0_n\}^*$ $\delta_z = 0_{n_z}{}^\spadesuit$ or $\delta_z > 0_{n_z}{}^\dagger$

Table 5

† $\Gamma\left(\bar{\delta}_z\right) = 1 - \beta_z$ for all $\bar{\delta}_z$ such that $\delta_z \geq \bar{\delta}_z \geq 0_{n_z}$.
$^\spadesuit$ $\langle x_o, x_z, y, z \rangle$ lies on the lower frontier of T_2.
* $\Theta\left(\bar{\delta}\right) = \beta_y - 1$ for all $\bar{\delta}$ such that $\delta \geq \bar{\delta} \geq 0_n$.
$^\clubsuit$ $\langle x_o, x_z, y, z \rangle$ is a strictly efficient point of sub-technology T_1.
When $\spadesuit$ and $\clubsuit$ are simultaneously true then $\langle x_o, x_z, y, z \rangle$ is a strictly efficient point of technology T.

The values that θ^* and γ^* can independently take are listed below. The images of function Θ and Γ defined in (24) and (25) are employed to infer the corresponding values of δ^* and δ_z^*.

- Row 1: $\theta^* \in (0, \theta^O)$, which implies $\delta^* \in L_\delta^{BP} \setminus \left(I_\delta^{BP} \cup \{0_n\}\right)$.
- Row 2: $\theta^* = 0$ but $\theta^* \neq \theta^O$, which implies $\delta^* \in I_\delta^{BP} \setminus \{0_n\}$.

- Row 3: $\theta^* = \theta^O$ but $\theta^* \neq 0$, which implies $\delta^* = 0$ or $\delta^* \in L_\delta^{BP} \setminus \left(I_\delta^{BP} \cup \{0_n\}\right)$.[46]

- Row 4: $\theta^* = \theta^O = 0$, which implies $\delta^* = 0$ or $\delta^* \in I_\delta^{BP} \setminus \{0_n\}$.[47]

- Column 1: $\gamma^* \in (\gamma^O, 1)$, which implies $\delta_z^* > 0_{n_z}$.

- Column 2: $\gamma^* = \gamma^O > 0$, which implies $\delta_z^* \geq 0_{n_z}$.[48]

- Column 3: $\gamma^* = \gamma^O = 0$, which implies $\delta_z^* \geq 0_{n_z}$.[49]

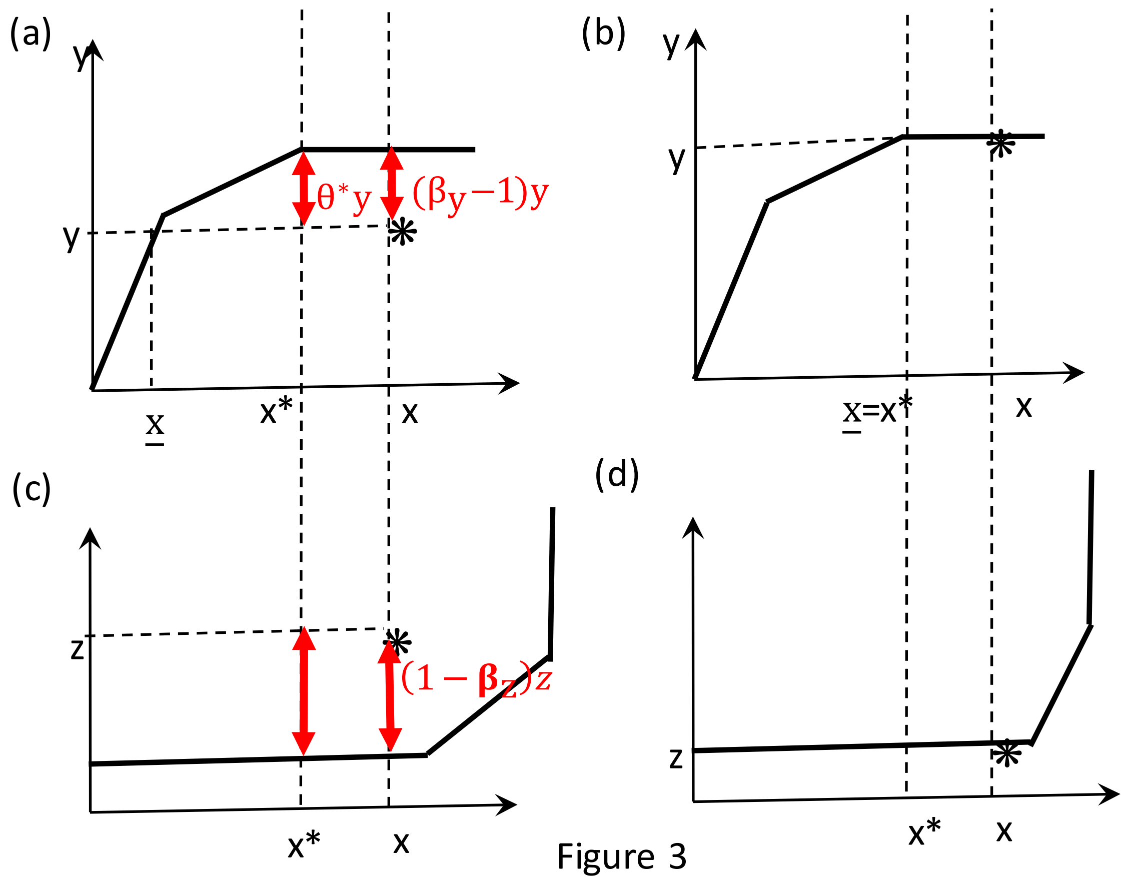

Figure 3

Recall that, under the BP approach, θ^O is the output-based index of productive efficiency improvement θ^O, while γ^O is the output-based index of environmental efficiency improvement γ^O. Note that, Columns 2 and 3 imply that whenever $\gamma^* - \gamma^O$ is positive, δ_z^* is positive. As seen in panels (c) and (d) of Figure 3, in general, it is possible that δ_z^* is positive, while $\gamma^* - \gamma^O$ is zero. But for a CEA-type dataset where there is a near perfect correlation between heat input and emission generation, under the BP approach, a positive δ_z^* will also imply that $\gamma^* - \gamma^O$ is positive.

[46]See panel (a) of Figure 3 for an example of a case where $\delta^* \in L_\delta^{BP} \setminus \left(I_\delta^{BP} \cup \{0_n\}\right)$.
[47]See panel (b) of Figure 3 for an example of a case where $\delta^* \in I_\delta^{BP} \setminus \{0_n\}$.
[48]See panel (c) of Figure 3 for an example where $\delta_z^* > 0_{n_z}$.
[49]See panel (d) of Figure 3 for an example where $\delta_z^* > 0_{n_z}$.

7.3 Optimal efficiency improvements in the India coal-based thermal power sector.

By categorising plants on the basis of the locations of their respective solutions to Problem (10) in Table 5, a thorough qualitative and quantitative analyses of the efficiency improvements recommended by the problem can be conducted. In this regards, it is helpful to first group plants based on the output-based measure of productive efficiency defined as

$$b_y := \frac{1}{1+\theta^O},$$

which is indicative of extent of technical efficiency relative to sub-technology T_1. This is because, as will be seen below, the nature and extent of optimal efficiency improvements in inputs and outputs are highly correlated to b_y.[50] We group all the coal-based power plants in our dataset into: (1) high performers, where b_y ranges between 0.95 to 1; (2) moderate performers, where b_y ranges between 0.795 to 0.95; and (3) low performers, where b_y ranges between 0 to 0.795.

To improve focus on studying the trade-offs between efficiency improvements in production of good and bad outputs, in our empirical analysis, we restrict ourselves to cases where zero weights are assigned to all inputs.[51] Conclusion of Theorem 7 for a BPT then implies that the optimum recommends no change in the usage of non-emission causing inputs. On the other hand, the optimum may recommend positive reductions in the usage of coal even when a zero weight is assigned to it. We first consider the the case where equal weights are assigned to efficiency improvements in both good and bad output production. We then study the consequences of increasing the weight on the efficiency improvement in generation of bad output.

7.3.1 Equal weights on θ and γ and zero weight on δ_z.

Optimal efficiency improvements for high performers.

Since the value of productive efficiency b_y is very high for high performers, they lie on or very close to the weakly efficient frontier of sub-technology T_1. Hence, Table 6 shows that the optimum of Problem 10 for these plants does not involve significant efficiency improvements in the usage of heat input and, hence, in the generation of emission (average values of δ_z^* and $\gamma^* - \gamma^O$ are both around 2%) and, in fact, it implies no efficiency improvement in the production of electricity (θ^* is zero for more than ninety-percent of these plants).

In particular, the optima of four of the high performers are located in Row 4 and Column 2 of Table 5. The former implies that $\theta^* = \theta^O = 0$ for these plants, and hence they are already

[50] The descriptive statistics of b_y are provided in Table 1.

[51] See Section 4.2.

operating on the weakly efficient frontier of T_1. The latter implies that $\gamma^* = \gamma^O > 0$. Thus, they are operating above the lower frontier of sub-technology T_2. It turns out that $\delta_z^* = 0$ for these plants, *i.e.,* the optimum does not recommend any change in usage of the heat input. Thus, the recommendations of both the graph and output-based measures of efficiency improvements coincide for these plants.

Table 6: Results with $w_y = w_z = 0.5$

		Performance categories		
		high	**moderate**	**low**
Solution categories from Table 5	(4,2)	4	0	0
	(2,1)	9	17	0
	(1,1)	1	2	2
	(3,2)	0	1	10
	(3,3)	0	0	1
θ*	Max	0.007	0.256	2.712
	Min	0	0	0.247
	Avg	0	0.018	0.716
δ_z*	Max	0.072	0.321	0.242
	Min	0	0	0
	Avg	0.024	0.184	0.022
$100(\gamma^*-\gamma^O)/\gamma^*$	Max	37.277	84.473	68.144
	Min	0	0	0
	Avg	15.190	56.162	7.893
$\gamma^*-\gamma^O$	Max	0.064	0.292	0.218
	Min	0	0	0
	Avg	0.022	0.165	0.019

The optimum for nine of the high performers is located in Row 2 and Column 1 of Table 5. This implies that optimal efficiency improvement in the heat input is not zero ($\delta_z^* > 0$) for each of these plants and leads to a point on the isoquant corresponding to the initial level y of electricity. Hence the optimal efficiency improvement in electricity θ^* is zero. At the same time the proportional reduction in the heat input usage δ_z^* contributes to a positive but small proportional reduction in emission: $\gamma^* - \gamma^O > 0$.

Optimal efficiency improvements for moderate performers.

As seen in Table 6, the optima of a majority of the moderate performers or plants that are placed farther away from the weakly efficient frontier of sub-technology T_1 are located in Row 2 and Column 1 of Table 5. Hence, the optimum continues to recommend (as in the case of majority of the high performers) reductions in the usage of the coal input with no change in electricity generation. But this time, the recommended reductions in the usage of the heat input are significant (δ_z^* can be as high as 32% with an average value of 18%) , which also imply significant reductions in CO_2 emission ($\gamma^* - \gamma^O$ can be as high as 29% with an average of 16.5%). In fact, plants under this category show maximum potential for reducing CO_2 emission when equal weights are assigned to good and bad outputs. A large part of the total proportional reduction in emission is attributed to reduction in usage of heat input: The share of $\gamma^* - \gamma^O$ in γ^* can be as high as 84% with an average value of 56%.

Optimal efficiency improvements for low performers.

These are plants that are located farthest from the weakly efficient frontier of sub-technology T_1. Thus, the extent to which electricity can potentially be increased with the existing levels of usage of heat input is greatest for these plants. The optimum of Problem 10 recommends that these plants focus exclusively on tapping this huge potential to increase electricity generation with their existing levels of usage of heat input. This implies that, for these plants, the gain from increase in electricity generation keeping heat input unchanged outweighs the gains from emission reduction that could have been achieved by reducing the heat input. Thus we find that a majority of these plants are located in Row 3 and Column 2 of Table 5, with $\delta_z^* = 0$ and $\theta^* = \theta^O$

Plants with interior solutions to Problem 10.

Table 6 shows that even under a weighting scheme that gives equal weightage to the good and bad outputs, for a majority of plants, the optimum of Problem 10 yields corner solutions: either it recommends efficiency improvement only along the heat input and emission directions with no increase in electricity generation (this is the case of high and moderate performers) or it recommends huge efficiency improvement along the electricity direction with no reduction in the heat input (as in the case of the low performers). In the former case, the gain in the objective function from increase in emission reduction achieved by reducing heat input (when moving along the lower frontier of T_2) must outweigh the loss from reduction in electricity generation due to the reduction in the heat input (when moving along the efficient frontier of T_1); while in the latter case, the reverse must be true.

There are only five plants in our data set for which Problem 10 yields interior solutions when equal and exhaustive weights are assigned to the good and bad outputs, *i.e.,* for these plants $\delta_z^* > 0$, $\theta^* > 0$, and $\gamma^* > 0$. These plants are positioned in Row (1) and Column (1) of Table 5.

7.4 Effects of varying weights on θ and γ.

Table 7 summarises the results when the weight on proportionate reduction in emission is increased gradually from zero to one-third, half, two-thirds, and finally to one. As we move to weighting schemes that attach more and more weight to the bad output CO_2 and less and less weight to the good output electricity, we find that the number of plants for which the optimum recommends minimising usage of the heat input and operating on the isoquants corresponding to their existing levels of electricity generation (*i.e.,* the number of plants for which the optimum lies in Row 2 of Table 5) increases.

Table 7: Result under different weighting schemes

		w_y= 2/3, w_z= 1/3	w_y= 1/2, w_z= 1/2	w_y= 1/3, w_z= 2/3	w_y= 0, w_z= 1
Solution categories from Table 5	(4,2)	4	4	4	4
	(2,1)	0	26	38	43
	(1,1)	0	5	3	0
	(3,2)	39	11	2	0
	(3,3)	1	1	0	0
	(3,1)	3	0	0	0
γ^*-γ^o	Max	0.025	0.292	0.520	0.685
	Min	0	0	0	0
	Avg	0.001	0.082	0.161	0.195

When full weight is assigned to emission reduction, then optima of all plants are located in Row 2 of Table 5. Hence, the maximum possible proportional reductions in the emission of CO_2 also increases as the weight on γ increases. It also turns out that, while the low performers showed maximum potential for electricity-expansion when equal weights were assigned to efficiency improvements in emission and electricity directions, they switch over to becoming the plants with the highest potential for proportionate reductions in heat-usage and hence CO_2 emission when the weight on γ increases by more than a half.

8 Conclusions.

The relations between four influential approaches to modelling emission-generating technologies and their implications for designing technical efficiency improvements for inefficient production units are explored in this paper. The theoretical analysis is combined with empirical applications that employ data on the Indian coal-based thermal power sector. While, under constant and non-increasing returns, the by-production approach satisfies joint-disposability of emission-causing inputs and the good and the bad outputs, in the class of DEA technologies, it stands unique in comparison to the weakly disposable, jointly disposable, and input approach-based technologies. The DEA-versions of the latter three classes of technologies satisfy a nesting relation and share common strictly efficient frontiers. This also implies that the efficiency improvements in inputs and outputs recommended by these three technology-modelling approaches are the same and differ from those recommended by the by-production approach.

The work highlights the importance of assessing technology modelling approaches on the basis of the extent to which they can successfully capture the production relations underlying the true data generating processes (DGPs). In this regards the CEA dataset employed in this paper provides a unique test, where the DGP for CO_2 is a deterministic linear formula involving constant emission factors of coal and oil, which results in an almost linear relation between CO_2 generation and an aggregate fossil-fuel input measured in heat units. Our empirical findings seem to indicate that the by-production approach is more successful in capturing these features of the DGP. This does not seem surprising as, under this approach, technology is modelled to incorporate multiple-production relations, and thus provides the needed flexibility to capture the relation between emission and emission-causing inputs seen in the DGP.

Under the by-production approach there is a trade-off between efficiency improvements in good and bad outputs: While the former are non-positively related to efficiency improvements in inputs, the latter are non-negatively related to efficiency improvements in emission-causing inputs. All possible configurations of optimal efficiency improvements in inputs and good and bad outputs under the by-production approach are tabulated in Table 5. This table is then used to analyse differences in optimal configurations of efficiency improvements for Indian coal-fired thermal plants power plants.

With equal weights assigned to good and bad outputs, we find that, for plants with high and moderate output-based efficiency, the graph-based efficiency measure recommends focussing exclusively on improving efficiency in the usage of the fossil-fuel input and generation of emission rather than on expanding the electrical output. This kind of efficiency improvement is basically *thermodynamic* in nature.[52] For plants with low output-based efficiency, the graph-based optimum recommends focusing exclusively on improving efficiency in generating electricity with existing levels of usage of the aggregate fossil fuel. These recommendations reflect costs and benefits of reduction in the use of the fossil fuel due to the associated trade-off in efficiency improvements in good and bad output production.

As we move to weighting schemes that attach more and more weight to mitigation of the emission, the number of plants for which the optimum recommends minimising usage of the fossil fuel and operating on the isoquants corresponding to their existing levels of electricity generation increases. On the other hand, as usage of the fossil fuel reduces, the maximum possible reduction in emission generation increases.

The computation of the weighted graph index of efficiency improvements yields an estimate of mitigation of emission possible with (at least, theoretically) no opportunity costs (in terms of losses in electricity generation) due to factors such as improvements in thermodynamic efficiency. When full weight is assigned to emission generation, technical efficiency improvements imply an estimate of up 14% reduction in usage of fossil fuel and 23% reduction in total emission of which around 55% can be attributed solely to improvements in thermodynamic efficiency (*i.e.,* reduction in the usage of fossil fuel) and the remaining can be attributed to improvements in

[52]Baumgärtner and Arons (2003) illustrate cases where thermodynamic inefficiencies can cause producing units to employ more than the minimal amounts of fossil-fuels required to produce given amounts of industrial outputs. This would imply generation of more than the minimal amounts of emissions while producing such outputs. They highlight that thermodynamic inefficiencies are rampant in real life. Removal of such inefficiencies imply minimising usage of fossil fuels and hence reductions in emission generation with no compromise (*i.e.,* no change) in electricity generation.

output-based efficiency (this is the potential efficiency improvements in emission generation if aggregate fossil-fuel usage was unchanged).[53] Based on Table 5 of all possible solutions, this paper identifies moderate and low performing plants that operate far below the efficient frontier of the intended-production technology T_1, a sub-technology of the by-production technology. Together these plants can contribute to 74% of the total reduction in emission and 92% of the reduction in emission attributable to efficient reductions in fossil-fuel usage (*i.e.,* improvements in thermodynamic efficiency). These should hence be the targeted units of any policy seeking to realise opportunities to mitigate CO_2 emissions with least opportunity costs.

REFERENCES

Baumgärtner, S., and J. de Swaan Arons (2003), "Necessity and Inefficiency in the Generation of Waste," *Journal of Industrial Ecology*, 7: 113–123.

Baumol, W. J., and W. E. Oates (1975, 1988), *The Theory of Environmental Policy*, 1st and 2nd editions, Cambridge University Press.

Behera, S.K., A.P. Dash and J. A. Farooquie (2010), "Performance Analysis of Coal Fired Power Plants in India," *Proceedings of the 2010 International Conference on Industrial Engineering and Operations Management Dhaka*, Bangladesh.

Boyd, G. A., and J. D. McClelland (1999), "The Impact of Environmental Constraints on Productivity Improvement in Integrated Paper Plants," *Journal of Environmental and Economic Management*, 38: 121–142.

Coggins, J. S., and J. R. Swinton (1996), "The Price of Pollution: A Dual Approach to Valuing SO_2 Allowances," *Journal of Environmental Economics and Management*, 30: 58–72.

Cropper, M. L., and W. E. Oates (1992), "Environmental Economics: A Survey," *Journal of Economic Literature*, 30: 675–740.

Debreu, G. (1959), *Theory of value*, New York, Wiley

Färe, R., S. Grosskopf, C. A. K. Lovell, and C. Pasurka (1989), "Multilateral Productivity Comparisons When Some Outputs are Undesirable: A Nonparametric Approach," *The Review of Economics and Statistics*, 71: 90–98.

Färe, R., and C.A.K. Lovell (1978) "Measuring the technical efficiency of production," *Journal of Economic Theory* 19, 150–162.

Färe, R, S. Grosskopf, and C. A. K. Lovell (1985), *The Measurement of Efficiency of Production*, Boston: Kluwer-Nijhoff.

Førsund, F. (1998), "Pollution Modeling and Multiple-Output Production Theory," Discussion Paper #D-37/1998, Department of Economics and Social Sciences, Agricultural University

[53]The output-based efficiency improvement, in contrast to thermodynamic efficiency improvement, may involve strategies such as switching to a more clean mix of fossil-fuel types, while holding the aggregate fossil-fuel usage (measured in heat units) fixed.

of Norway.

Førsund, F. (2009), "Good Modelling of Bad Outputs: Pollution and Multiple-Output Production," *International Review of Environmental and Resource Economics*, 3, 1–38.

Førsund, F. (2017), "Multi-equation Modeling of Desirable and Undesirable Outputs Satisfying the Material Balance," *Empirical Economics*, online.

Frisch, R. (1965), *Theory of Production*, Dordrecht, D. Reidel Publishing Company.

Kumar, R., A. K. Sharma, and P. C. Tewari (2015), "Cost analysis of a coal-fired power plant using the NPV method," *Journal of Industrial Engineering International*, DOI 10.1007/s40092-015-0116-8

Mas-Colell, A., M. D. Whinston, and J. R. Green (1995) *Microeconomic theory*, Chp. 5. Oxford University Press, New York

Ministry of Power (2013-14, 2014-15), *Review of Performance of Thermal Power Plants in India*, Central Electricity Authority, Ministry of Power, Government of India.

Murty, M. N., and S. Kumar (2002), "Measuring Cost of Environmentally Sustainable Industrial Development in India: A Distance Function Approach," *Environment and Development Economics*, 7: 467–86.

Murty, M. N., and S. Kumar (2003), "Win-Win Opportunities and Environmental Regulation: Testing of Porter Hypothesis for Indian Manufacturing Industries," *Journal of Environmental Management*, 67: 139–44.

Murty, S (2015), "On the Properties of an Emission-Generating Technology and its Parametric Representation," *Economic Theory*. 60, 243–282.

Murty, S, and R. Nagpal (2017), "Measuring Output-Based Technical Efficiency of Indian Coal-Based Thermal Power Plants: A By-production Approach," Discussion Paper 18-07, Centre for International Trade and Development, School of International Studies, Jawaharlal Nehru University. $https://jnu.ac.in/sites/default/files/citd/CITD-DP-Murty\%26NagpalOuputBasedEfficiency2018.pdf$

Murty, S, and R. R. Russell (2002), "On Modeling Pollution-Generating Technologies," Department of Economics, University of California, Riverside, Discussion Papers Series, No. 02-14.

Murty, S., R. R. Russell, and S. B. Levkoff (2012), "On Modeling Pollution-Generating Technologies," *Journal of Environmental Economics and Management*, 64, 117–135.

Murty, S., and R. R. Russell (2017) "Bad outputs," Discussion Papers in Economics No. 17-06, Centre for International Trade and Development, School of International Studies, Jawaharlal Nehru University, India.

Murty, S., and R. R. Russell (2018), "Modeling Emission-Generating Technologies: Reconciliation of Axiomatic and By-Production Approaches," *Empirical Economics*, 54: 7–30 (doi version (2016)).

Ray, S. C., K. Mukherjee, and A. Venkatesh (2018), "Nonparametric Measures of Efficiency in the Presence of Undesirable Outputs: A By-production Approach with Weak Disposability," *Empirical Economics*, 54: 31–65 (doi version (2016)).

Sahoo, N. R., Pratap K.J. Mohapatra, B. K. Sahoo, and B. Mahanty (2017) "Rationality of Energy Efficiency Improvement Targets under the PAT Scheme in India – A Case of Thermal Power Plants," *Energy Economics* doi:10.1016/j.eneco.2017.06.004

Serra, T., R. G. Chambers, and A. O. Lansink (2016), "Measuring Technical and Environmental Efficiency in a State-Contingent Techology," *European Journal of Operational Research*, 236: 706–717.

Shephard, R. W. (1953), *Cost and production functions*, Princeton University Press, Princeton.

Hailu A., and T. S. Veeman (2000), "Environmentally Sensitive Productivity Analysis of the Canadian Pulp and Paper Industry, 1959–1994: An Input Distance Function Approach," *Journal of Environmental Economics and Management*, 40, 251-274.

Tong D., Q. Zhang, S. J. Davis, F. Liu, B. Zheng, G. Geng, T. Xue, M. Li, C. Hong, Z. Lu, D. G. Streets, D. Guan and K. He (2018),"Targeted emission reductions from global super-polluting power plant units," *Nature Sustainability*, 1, 59–68

Zhou P., B. Ang, and K-L Poh (2008), "A Survey of Data Envelopment Analysis in Energy and Environmental Studies," *European Journal of Operational Research*, 189: 1–18.

APPENDIX

Proof. of Theorem 4

(i) Given the data matrices X, Y, and Z, if $v = \langle x, y, z\rangle$ is a strictly efficient point of technology T^I, then there exists intensity vector λ such that the following equalities hold: $\lambda^T X = x$, $\lambda^T Y = y$, and $\lambda^T Z = z$.[54] Hence, $v = \langle x, y, z\rangle = \left\langle \lambda^T X,\ \lambda^T Y,\ \lambda^T Z\right\rangle$. The constructs (4) and (5) also imply that $v = \langle x, y, z\rangle = \left\langle \lambda^T X,\ \lambda^T Y,\ \lambda^T Z\right\rangle \in T^\kappa$ for $\kappa = WD, JD$. Hence, (7) implies that it is also a strictly efficient point of T^κ for $\kappa = WD, JD$.

If $v = \langle x, y, z\rangle = \left\langle \lambda^T X,\ \lambda^T Y,\ \lambda^T Z\right\rangle \in T^\kappa$ is a strictly efficient point of T^κ for $\kappa = WD, JD$, then v also satisfies (6). Hence, $v \in T^I$. If v is not a strictly efficient point of T^I, then solving problem (10) with $T = T^I$ yields a production vector $v^* = \langle x_o^*, x_z^*, y^*, z^*\rangle$ that is a strictly efficient point of T^I with $x_o^* \leq x_o$, $x_z^* \leq x_z$, $y^* \geq y$, and $z^* \leq z$, with at least one of the weak inequalities holding as a strict inequality. But then, as argues in the first part of the proof, v^* is also a strictly efficient point of T^κ, which contradicts v being a strictly efficient point of T^κ.

(ii) If $\langle y, z\rangle$ is a strictly efficient point of $P^I(x)$, then there exists $\lambda \geq 0_U$ such that $\lambda^\top X \leq x^\top$, $\lambda^\top Y = y$, $\lambda^\top z = z$. (9) implies that $\langle y, z\rangle$ is in $P^{WD}(x)$. From (9) it is also clear that

[54] In this proof, we will maintain that, if constant returns to scale is true then intensity vectors satisfy $\lambda \geq 0_U$. If non-increasing returns to scale is true, then intensity vectors satisfy $\lambda \in [0, 1]^U$.

$P^{WD}(x) \subset P^I(x)$. Hence, $\langle y, z\rangle$ is also a strictly efficient point of $P^{WD}(x)$.
Suppose $\langle y, z\rangle$ is a strictly efficient point of $P^{WD}(x)$ but is not a strictly efficient point of $P^I(x)$. Then problem (11) yields solution $\langle \theta_I^O, \gamma_I^O\rangle$ for technology T^I such that (a) $\langle y + \theta_I^O y, z - \gamma_I^O z\rangle$ is a strictly efficient point of $P^I(x)$ and (b) $y + \theta_I^O y \geq y$ and $z - \gamma_I^O z \leq z$ with at least one of the two weak inequalities holding as a strict inequality. (a) implies that there exist $\lambda \geq 0_U$ such that $\lambda^\top Y = y + \theta_I^O y$, $\lambda^\top Z = z - \gamma_I^O z$, and $\lambda^\top X \leq x^\top$. Hence, (9) implies that $\langle y + \theta_I^O y, z - \gamma_I^O z\rangle \in P^{WD}(x)$. But (b) implies that this in contradiction of our assumption that $\langle y, z\rangle$ is a strictly efficient point of $P^{WD}(x)$. Hence, $\langle y, z\rangle$ is also a strictly efficient point of $P^I(x)$. ■

Proof. (Theorem 12) Suppose $\langle \bar{\delta}_o, \bar{\delta}_z\rangle$ and $\langle \hat{\delta}_o, \hat{\delta}_z\rangle$ are both in $[0,1]^n$ and $\langle \bar{\delta}_o, \bar{\delta}_z\rangle \geq \langle \hat{\delta}_o, \hat{\delta}_z\rangle$. Let $\hat{\theta}^\kappa = \Theta^\kappa\left(\hat{\delta}_o, \hat{\delta}_z\right)$, $\bar{\theta}^\kappa = \Theta^\kappa\left(\bar{\delta}_o, \bar{\delta}_z\right)$, $\hat{\gamma}^\kappa = \Gamma^\kappa\left(\hat{\delta}_o, \hat{\delta}_z\right)$, and $\bar{\gamma}^\kappa = \Gamma^\kappa\left(\bar{\delta}_o, \bar{\delta}_z\right)$ for $\kappa \in \{WD, JD, I, BP\}$. Then

$$\left\langle x_o - \left(\bar{\delta}_o \otimes x_o\right),\ x_z - \left(\bar{\delta}_z \otimes x_z\right),\ y + \bar{\theta}^\kappa y,\ z - \bar{\gamma}^\kappa z\right\rangle \in T^\kappa. \tag{26}$$

(i) For Suppose $\kappa \in \{WD, I\}$. Since T^κ satisfies free input disposability and

$$\left\langle x_o - \left(\hat{\delta}_o \otimes x_o\right),\ x_z - \left(\hat{\delta}_z \otimes x_z\right)\right\rangle \geq \left\langle x_o - \left(\bar{\delta}_o \otimes x_o\right),\ x_z - \left(\bar{\delta}_z \otimes x_z\right)\right\rangle, \tag{27}$$

we have

$$\left\langle x_o - \left(\hat{\delta}_o \otimes x_o\right),\ x_z - \left(\hat{\delta}_z \otimes x_z\right),\ y + \bar{\theta}^\kappa y,\ z - \bar{\gamma}^\kappa z\right\rangle \in T^\kappa.$$

Therefore, $\langle \bar{\theta}^\kappa, \bar{\gamma}^\kappa\rangle$ is in the constraint set of problem (15) when $\langle \delta_o, \delta_z\rangle = \langle \hat{\delta}_o, \hat{\delta}_z\rangle$.
Hence, $w^y\hat{\theta}^\kappa + w^z\hat{\gamma}^\kappa \geq w^y\bar{\theta}^\kappa + w^z\bar{\gamma}^\kappa$.
(ii) Suppose $\kappa = JD$ and $\bar{v} := \left\langle x_o - \left(\bar{\delta}_o \otimes x_o\right),\ x_z - \left(\bar{\delta}_z \otimes x_z\right),\ y + \bar{\theta}^{JD} y,\ z - \bar{\gamma}^{JD} z\right\rangle$ and $\hat{v} := \left\langle x_o - \left(\hat{\delta}_o \otimes x_o\right),\ x_z - \left(\hat{\delta}_z \otimes x_z\right),\ y + \hat{\theta}^{JD} y,\ z - \hat{\gamma}^{JD} z\right\rangle$ are strictly efficient points of T^{JD}. Proposition 4 implies that $\bar{v}$ and $\hat{v}$ are also efficient points of T^{WD}. Hence, $\Theta^{WD}\left(\bar{\delta}_o, \bar{\delta}_z\right) = \bar{\theta}^{JD}$ and $\Gamma^{WD}\left(\bar{\delta}_o, \bar{\delta}_z\right) = \bar{\gamma}^{JD}$. Similarly, $\Theta^{WD}\left(\hat{\delta}_o, \hat{\delta}_z\right) = \hat{\theta}^{JD}$ and $\Gamma^{WD}\left(\hat{\delta}_o, \hat{\delta}_z\right) = \hat{\gamma}^{JD}$. Hence, conclusion (i) of this proposition implies that $w^y\hat{\theta}^{JD} + w^z\hat{\gamma}^{JD} \geq w^y\bar{\theta}^{JD} + w^z\bar{\gamma}^{JD}$.
(iii) (a) Proving monotonicity of functions Θ^{BP} and Γ^{BP}:
(26) implies

$$\left\langle x_o - \left(\bar{\delta}_o \otimes x_o\right),\ x_z - \left(\bar{\delta}_z \otimes x_z\right),\ y + \bar{\theta}^{BP} y,\ z\right\rangle \in T_1.$$

Since T_1 satisfies free input disposability and (27) is true, we have

$$\left\langle x_o - \left(\hat{\delta}_o \otimes x_o\right),\ x_z - \left(\hat{\delta}_z \otimes x_z\right),\ y + \bar{\theta}^{BP} y,\ z\right\rangle \in T_1.$$

Therefore, $\bar{\theta}^{BP}$ is in the constraint set of problem (17) when $\langle\delta_o, \delta_z\rangle = \langle\hat{\delta}_o, \hat{\delta}_z\rangle$. Hence $\hat{\theta}^{BP} \geq \bar{\theta}^{BP}$.

Since $\bar{\delta} > \hat{\delta}$, we have $\bar{\delta}_z \geq \hat{\delta}_z$. Also $\hat{\gamma}^{BP} = \Gamma^{BP}\left(\hat{\delta}_o, \hat{\delta}_z\right) = \Gamma\left(\hat{\delta}_z\right)$ and $\bar{\gamma}^{BP} = \Gamma^{BP}\left(\bar{\delta}_o, \bar{\delta}_z\right) = \Gamma\left(\bar{\delta}_z\right)$. Given definition of function Γ in (18), we have:

$$\left\langle x_o,\ x_z - \left(\hat{\delta}_z \otimes x_z\right),\ y,\ z - \hat{\gamma}^{BP}\bar{z}\right\rangle \in CDH\left(T_2\right).$$

Since $CDH\left(T_2\right)$ satisfies the assumptions of costly disposability emission and emission-causing inputs, and

$$x_z - \left(\hat{\delta}_z \otimes x_z\right) \geq x_z - \left(\bar{\delta}_z \otimes x_z\right),$$

we have

$$\left\langle x_o,\ x_z - \left(\bar{\delta}_z \otimes x_z\right),\ y,\ z - \hat{\gamma}^{BP} z\right\rangle \in CDH\left(T_2\right).$$

Therefore $\hat{\gamma}^{BP}$ is in the constraint set of problem (18) when $\delta_z = \bar{\delta}_z$. Hence $\bar{\gamma}^{BP} \geq \hat{\gamma}^{BP}$.

(b) Proving concavity of $\Theta^{BP} = \Theta$ and $\Gamma^{BP} = \Gamma$:

We need to show that the hypographs of functions Θ and Γ are convex sets.[55] Let $\delta := \langle\delta_o, \delta_z\rangle$ and $\hat{\delta} := \langle\hat{\delta}_o, \hat{\delta}_z\rangle$ lie in Ω and define $\delta' = \alpha\delta + (1-\alpha)\hat{\delta}$, where $\alpha \in [0,1]$. Let $\Theta(\delta) = \theta$, $\Theta\left(\hat{\delta}\right) = \hat{\theta}$, and $\Theta(\delta') = \theta'$. Thus, $\langle\delta, \theta\rangle$, $\langle\hat{\delta}, \hat{\theta}\rangle$ lie in the hypograph of Θ. To show that the hypograph of Θ is a convex set, we need to show that $\langle\delta', \theta^*\rangle$ is also in the hypograph of Θ, where $\theta^* = \alpha\theta + (1-\alpha)\hat{\theta}$. That is, we need to show that

$$\theta^* \leq \Theta(\delta') \equiv \theta'.$$

The definition of function Θ implies that $v = \langle x + \delta \otimes x,\ y + \theta y,\ z\rangle \in T_1$, $\hat{v} = \langle x + \hat{\delta} \otimes x,\ y + \hat{\theta} y,\ z\rangle \in T_1$, and $v' = \langle x + \delta' x,\ y + \theta' y,\ z\rangle \in T_1$. Since T_1 is a convex set, we have $\alpha v + (1-\alpha)\hat{v} \in T_1$. But $\alpha v + (1-\alpha)\hat{v} = \langle x + \delta' \otimes x,\ y + \theta^* y,\ z\rangle \in T_1$. Thus, θ^* is in the constraint set of problem (17) when the proportional change in inputs is given by the vector δ'. Hence, the definition of Θ implies that $\theta^* \leq \Theta(\delta')$. Hence, $\langle\delta', \theta^*\rangle$ is in the hypograph of Θ.
The proof of concavity of function Γ proceeds in an exactly similar manner. ■

Proof. (Theorem 7)
Let $v^* := \langle x_o^*, x_z^*, y^*, z^*\rangle = \langle x_o - (\delta_o^* \otimes x_o),\ x_z - (\delta_z^* \otimes x_z), y + \theta^* y,\ z - \gamma^* z\rangle$. Then $v^* \in T$. Hence $x^* = \langle x_o^*, x_z^*\rangle \in L(y,z) := \{x' \in \mathbf{R}_+^n \mid \langle x', y, z\rangle \in T\}$.

[55] The hypograph of a function $f : \mathbf{R}^n \longrightarrow \mathbf{R}$ with image $y = f(x)$ is defined as the set $\{\langle x, y\rangle \in \mathbf{R}^{n+1} \mid y \leq f(x)\}$.

(i) Since $w_i^x = 0$, the objective function evaluated at the optimum is

$$w_i^x \delta_i^* + \sum_{j \neq i} w_j^x \delta_j^* + w^y \theta^* + w^z \gamma^* = w_i^x \bar{\delta}_i + \sum_{j \neq i} w_j^x \delta_j^* + w^y \theta^* + w^z \gamma^*, \quad \forall\ \bar{\delta}_i \in [0, 1].$$

In particular, we show below that choosing $\bar{\delta}_i = 0$ is technologically feasible. This would imply that $\bar{s} = \left\langle \bar{\delta}_i = 0, \delta_{-i}^*, \theta^*, \gamma^* \right\rangle$ is also a solution to problem (10).

Since T satisfies free input disposability with respect to input i and $\langle x_i, x_{-i}^* \rangle \geq \langle x_i^*, x_{-i}^* \rangle$, we have $\langle x_i, x_{-i}^* \rangle \in L(y, z)$. Hence, starting from v, the vector of efficiency improvements $\left\langle \bar{\delta}_i = 0, \delta_{-i}^*, \theta^*, \gamma^* \right\rangle$ is technologically feasible as it leads to $\langle x_i, x_{-i}^*, y^*, z^* \rangle$, which is also in T.

(ii) Suppose $\delta_i^* > 0$. The latter implies that $x_i > x_i^*$. Since T satisfies free input disposability with respect to input i and $\langle x_i, x_{-i}^* \rangle > \langle x_i^*, x_{-i}^* \rangle$, we have $\langle x_i, x_{-i}^* \rangle \in L(y, z)$. Thus, $\lambda = 1$ is in the constraint set of the problem:

$$\lambda^* := \max \left\{ \lambda > 0 \;\middle|\; \left\langle x_i, \frac{x_{-i}^*}{\lambda} \right\rangle \in L(y, z) \right\}, \tag{28}$$

so that $\lambda^* \geq 1$. Since Ψ_T is increasing in input i, we have

$$y = \Psi_T \left(x_i^*, x_{-i}^*, z \right) \ < \ \Psi_T \left(x_i, x_{-i}^*, z^* \right).$$

At the solution of problem (28), we have

$$y = \Psi_T \left(x_i, \frac{x_{-i}^*}{\lambda^*}, z^* \right).$$

Thus we have

$$y = \Psi_T \left(x_i, \frac{x_{-i}^*}{\lambda^*}, z^* \right) \ < \ \Psi_T \left(x_i, x_{-i}^*, z^* \right) \implies \lambda^* > 1 \text{ and } \frac{x_{-i}^*}{\lambda^*} < x_{-i}^*.$$

Then $\langle x_i, \bar{x}_{-i}, y^*, z^* \rangle \in T$, where $\bar{x}_{-i} := \frac{x_{-i}^*}{\lambda^*} < x_{-i}^*$. This implies $\delta_j^* = \frac{x_j - x_j^*}{x_j} \leq \frac{x_j - \bar{x}_j}{x_j} =: \bar{\delta}_j$ for all $j \neq i$ with the weak inequality holding strictly for at least one $j \neq i$. Then the vector of efficiency improvements $\bar{s} = \langle \delta_i^*, \bar{\delta}_{-i}, \theta^*, \gamma^* \rangle$ is in the constraint set of problem (10), as it leads to production vector $\langle x_i^*, \bar{x}_{-i}, y^*, z^* \rangle \in T$. We also have

$$w_i^x \delta_i^* + \sum_{j \neq i} w_j^x \bar{\delta}_j + w^y \theta^* + w^z \gamma^* > \sum_{j=1}^{n} w_j^x \delta_j^* + w^y \theta^* + w^z \gamma^*.$$

This is a contradiction to $s^* = \langle \delta_o^*, \delta_z^*, \theta^*, \gamma^* \rangle$ being a solution of problem (10). Hence, $\delta_i^* = 0$. ■

Proof. (Theorem 14)

Image of Θ*:* Let $\theta^* := \Theta(\delta_o, \delta_z)$.

- Case: $\langle \delta_o, \delta_z \rangle = 0_n$: This implies $\langle x_o - (\delta_o \otimes x_o), \ x_z - (\delta_z \otimes x_z) \rangle = \langle x_o, x_z \rangle$. Since $v = \langle x_o, x_z, y, z \rangle \in T_1$, we have $\langle \delta_o, \delta_z \rangle = 0_n \in L_\delta^{BP}$. If $\langle \delta_o, \delta_z \rangle = 0_n$ and θ^* solves problem (17), *i.e.,* θ^* solves problem

$$\max \{\theta \in \mathbf{R} \mid \langle x_o, x_z, (1+\theta)y, z \rangle \in T_1\}. \tag{29}$$

  A comparison with Problem (**??**) implies that $\boldsymbol{\beta}_y(x_o, x_z, y, z) = 1 + \theta^*$. Hence, $\theta^* = \theta^O$.

- Case: $\langle \delta_o, \delta_z \rangle \in I_\delta^{BP} \setminus \{0\}$: This implies $\langle x_o - (\delta_o \otimes x_o), \ x_z - (\delta_z \otimes x_z) \rangle \in I^{BP}(y)$. Hence, $y = F(x_o - (\delta_o \otimes x_o), \ x_z - (\delta_z \otimes x_z))$. Remark 13 then implies $(1+\theta^*)y = y$. Hence, $\theta^* = 0$.

- Case: $\langle \delta_o, \delta_z \rangle \in L_\delta^{BP} \setminus \left(I_\delta^{BP} \cup \{0_n\}\right)$: Combined with Remark 13, this case implies that

$$y < F(\bar{x}_o - (\delta_o \otimes \bar{x}_o), \ \bar{x}_z - (\delta_z \otimes \bar{x}_z)) = (1+\theta^*)y =: y^*.$$

  Hence, $\theta^* > 0$.
  Since $L_\delta^{BP} \subseteq \mathfrak{Q}$, we have $\langle \delta_o, \delta_z \rangle \in \mathfrak{Q} = [0,1]^n$. Since $\langle \delta_o, \delta_z \rangle \neq 0_n$, this implies that $\langle \delta_o, \delta_z \rangle > 0_n$. Since Θ is a non-increasing function (see Theorem 12) and $\Theta(0_n) = \theta^O$ and $\langle \delta_o, \delta_z \rangle > 0_n$, we have $\theta^* \leq \theta^O$. Thus, $\theta^* \in (0, \theta^O]$.

- Case: $\langle \delta_o, \delta_z \rangle \in \mathfrak{Q} \setminus L_\delta^{BP}$: This implies that, if $\langle x_o - (\delta_o \otimes x_o), \ x_z - (\delta_z \otimes x_z) \rangle \in L(\bar{y})$, then $\bar{y} < y$. This and Remark 13 imply that $y^* := (1+\theta^*)y = F(x_o - (\delta_o \otimes x_o), \ x_z - (\delta_z \otimes x_z)) < y$. Hence, $\theta^* < 0$.
  Non-negativity of the intended output implies that there is lower bound on the value function Θ can take: $\Theta(\delta_o, \delta_z) \geq -1$.
  Thus, if $\delta \in \mathfrak{Q} \setminus L_\delta^{BP}$, we have $-1 \leq \theta^* < 0$.

Image of Γ*:* Let $\gamma^* = \Gamma(\delta_z)$.
Non-negativity of technologically feasible levels of the emission implies that $x_z - \gamma^* x_z \geq 0$. Hence, γ^* has to take values that are no-bigger than one.
The minimum level of emission under $CDH(T_2)$ corresponding to vector of emission causing inputs x_z is $\beta_z z$. Since set $CDH(T_2)$ satisfies costly disposability, $\beta_z z$ level of emission is feasible under set $CDH(T_2)$ for all emission-causing input vectors smaller that x_z. Hence, $\gamma = \gamma^O$ is a member of the constraint set of problem (18) for all $\delta_z \in \mathfrak{Q}_z = [0,1]^{n_z}$. Hence $\Gamma(\delta_z)$

cannot take values less than γ^O for all $\delta_z \in \mathcal{Q}_z = [0,1]^{n_z}$. Thus we conclude that $\Gamma(\delta_z) \in [\gamma^O, 1]$ whenever $\delta_z \in \mathcal{Q}_z$.

In particular, when $\delta_z = 0_{n_z}$, then the fact that γ^* solves problem (18) implies that it is the solution of problem: $\max\{\gamma \in \mathbf{R}_+ \mid \langle x_o, x_z, y, (1-\gamma)z\rangle \in CDH(T_2)\}$. This in turn implies that $\boldsymbol{\beta}_z(x_o, x_z, y, z) = 1 - \gamma^*$. Hence, $\gamma^* = \gamma^O$. $\blacksquare$